AN ABSTRACT OF THE THESIS OF

Paul Anthony Brotherton for the degree of Master of Science in Industrial Engineering. Presented on May 5, 1999. Title: The Effects of Broad Based Holistic Measurement on Student Engagement and Motivation in Educational Assessment.

Abstract Approved: _____
Kimberly D. Douglas

Educational institutions are exposed to a continuously changing environment. As in business, they experiment with their methods and techniques to improve their outputs. Performance assessment, both in education and business is a gray science. It is not just a question of evaluating performance. Performance is comprised of a number of components such as motivation, ability, organizational support, and rewards.

This study looked at the relationship between educational assessment and the performance components. An assessment tool called the Balanced Scorecard, which has seen great success in the world of business, was adapted for use in the classroom. The study utilized quasi-experimental design to compare the effects of the broad-based holistic measurement associated with a balanced scorecard, and a traditional grading structure in two topics-based college courses. The study found that motivation, individual equity, satisfaction, and student engagement were all significantly higher in the experimental group by comparison. This evidence suggests that by utilizing a broad base of performance measures, one can increase student motivation and engagement in the learning process.

THE EFFECTS OF BROAD BASED HOLISTIC MEASUREMENT ON STUDENT ENGAGEMENT AND MOTIVATION IN EDUCATIONAL ASSESSMENT

By

Paul Anthony Brotherton

A Thesis Submitted

to

Oregon State University

in Partial Fulfillment of
the requirements for the
degree of

Master of Science

Presented May 5, 1999
Commencement June 13, 1999

Master of Science thesis of Paul Anthony Brotherton presented on May 5, 1999.

APPROVED:

Redacted for Privacy

Major Professor, representing Industrial Engineering

Redacted for Privacy

Chair of Department of Industrial and Manufacturing Engineering

Redacted for Privacy

Dean of Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for Privacy

Paul Anthony Brotherton, Author

Acknowledgements

There is quite a list of people who helped me accomplish this research and finish my master's degree. To Dr. Kimberly Douglas, mentor, advisor, friend, and personal cheering section, this wouldn't have been possible without her. I thank her for making me read that first article about the Balanced Scorecard, forever changing the course of my education. She has made my education fun, and what would an education be without fun.

I want to thank all of the students who participated in this research, many of whom I now called friends. Their time, patience, and willing contributions have made this an incredible experience - one I would gladly do all over again.

I also want to thank Megan Schneider and Dr. John Shea. Their willingness to participate in this experiment made this study possible. They made sure I had what I needed to accomplish my objectives.

I want to thank Anita Lauhala and Ginnie Kosse. They came to my rescue just as I started ripping my hair out. They gave me two days of their valuable time in order to pull off the focus groups together.

Saving the best for last, I thank my family. They thought I had fallen from the face of the earth during this project. They have supported everything I have ever done or ever will do. They convinced me that there was nothing I couldn't do if I set my mind to it. Hey look, it's done!

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF APPENDIX FIGURES

LIST OF APPENDIX FIGURES (Continued)

The Effects of Broad Based Holistic Measurement on Student Engagement and Motivation in Educational Assessment

Chapter 1: Introduction

1.1 Background

Do you know what the meaning of an A is? Does it mean a job well done? Is it just a letter? Where does it come from? What value does it have? It seems that we have come to accept it within education without a lot of thought. Our society has used and come to treat grades within education as an incentive much the same way that money is used to reward work on the job.

Assessment within education is grounds for a great deal of debate. Numerous theories and plans to improve assessment methods have been proposed. The issue of student motivation is often neglected in assessment design (Reineke, 1998). The focus, instead, is on ranking and quantifying performance. In fact, the prevailing work that has been done in the area of assessment addresses how to improve methods to make them more efficient (Rowntree, 1987). This approach suggests that assessment is necessary, and the goal is to improve the assessment methods.

Another argument is for the complete elimination of grades within education, on the grounds that it reduces the joy of learning (Deming, 1994).

Deming argues that the joy of learning is an intrinsic motivation that is destroyed by the introduction of extrinsic motivators such as grades. Figure 1.1 was developed by Deming to illustrate the effects of societal forces on intrinsic motivation during a person's life.

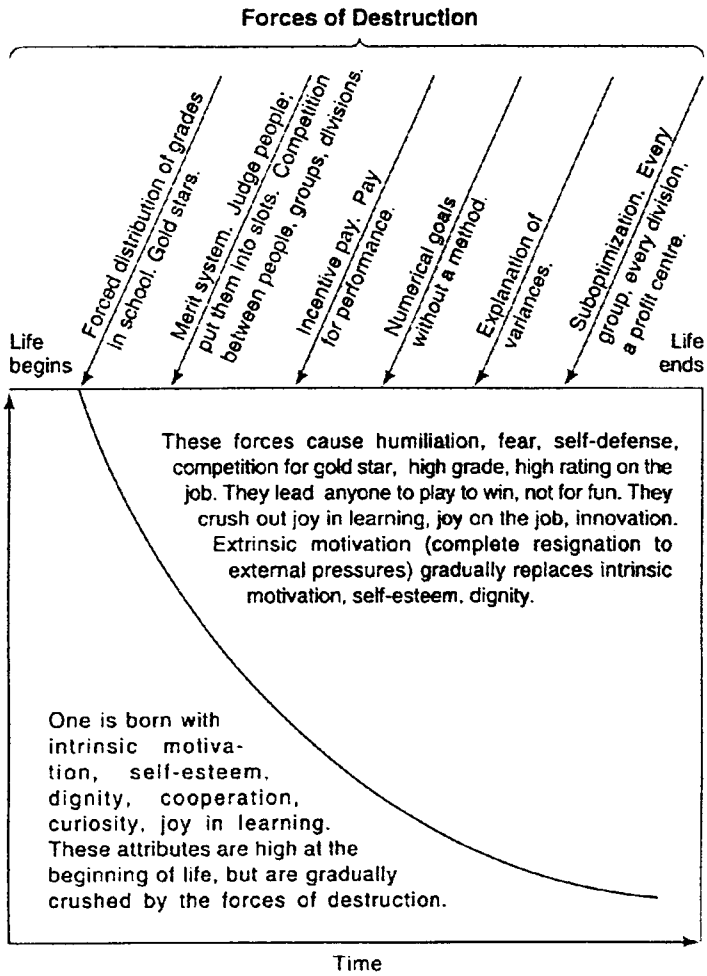


Figure 1.1: The effects of societal forces on intrinsic motivation of a person over time (Deming, 1994)

Kohn suggests that educators don't need to motivate students because they are internally driven to learn (1998). He also agrees there are striking parallels between the work environment and the educational environment. Yet if we do not grade students, how do we know how they are doing?

Is it possible that both arguments are right? At first glance they would seem to be mutually exclusive. Perhaps the true irony is the failure of both sides to recognize the importance of the other in the total assessment picture. Assessment impacts student motivation, and it is required in order to help guide and direct education. It is doubtful a company would be willing to hire graduates if they couldn't show technical competency through assessment as well as a high level of motivation. Therefore some sort of compromise must be reached between the two extreme arguments. The result would be an assessment method that limits the forces that slowly destroy intrinsic motivation, while at the same time providing valuable feedback to students, instructors, and other potential customers.

1.2 The Scope of This Research

The range of variables that are covered by the broad heading of educational assessment and motivation is extensive. The greatest challenge in this research was attempting to target a single phenomenon for study, which would provide useful information for the future. This was a difficult process and required multiple cycles and refinements before the final scope was defined.

The initial target was to study the construct of joy in learning, but the scale of this construct made the scope too large. From the initial target, several decisions were made to refine the issue into a more manageable research topic. The first was that a student focus was desired. The second was that the phenomenon of interest was an element of student attitude and not their observable performance.

The final decision was to study student satisfaction. Satisfaction fulfilled both of the established criteria. The other key characteristic of satisfaction was its importance in the workplace. There are several theories that point out the importance of satisfaction in the work place such as Fulfillment Theory, Discrepancy Theory, Equity Theory, and Two Factor Theory (Lawler, 1994). Satisfaction has profound impacts on motivation, productivity, and desire to perform. These theories show that satisfaction is not an independent variable, but part of a performance cycle.

1.2.1 Research Objective

The purpose of this research is to apply a broad based or balanced assessment mechanism in an educational environment and discover its effect on student motivation, and whether it could potentially create a positive escalation in the student performance cycle. The successful outcome of this research showed the importance of targeting motivation in the design of assessment system. Motivation is a common missing link in many assessment strategies.

1.2.2 Conceptual Framework

This research relied upon previous work done in the areas of assessment and motivation, across both the business and education environment. The most important factor in this research was that performance is a dynamic process, and that past experiences will effect future behaviors. Figure 1.2 details the ongoing dynamic between motivation, performance, and satisfaction. Assessment is a fundamental part of this dynamic and is part of the organizational support structure.

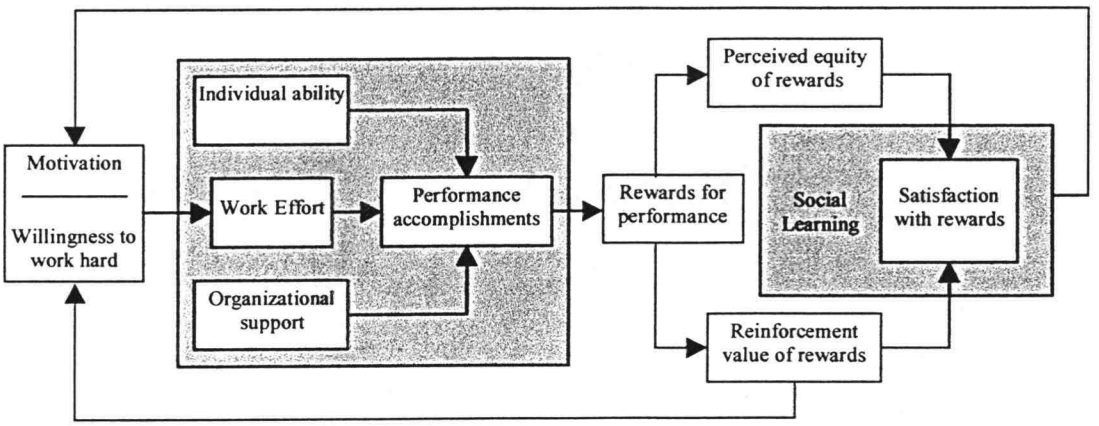


Figure 1.2: Managing motivational dynamics: an integrated approach to individual motivation to work. (Schermerborn, 1993)

According to Schermerborn, motivation is the foundation of the performance cycle. It will determine an individual's desire to perform, and is a function of expected outcomes (1993). The assessment links an individual's expectations and their externally perceived performance. Therefore, it is important that assessment methods take into consideration this dynamic. Satisfaction from

one experience becomes an input component of future iterations of the performance cycle.

A popular assessment method that has quickly gained support in the business world is the balanced scorecard. The primary reason that it has been adopted so quickly is its balanced or holistic view of an organization. It utilizes a broad base of performance measures that help to more clearly gauge organizational health. In the same way the balanced scorecard has helped business, it is the contention of the researcher that a scorecard that will more clearly gauge the performance of a student. The motivational dynamics model suggests that a better assessment of performance will improve satisfaction. Higher satisfaction becomes an input into motivation for the next cycle, thus creating positive escalation. A more detailed discussion of these constructs is located in the body of knowledge review.

This research focused on student motivation and in particular student satisfaction. In a dynamic cycle, no variable stands alone. Therefore, this research needed to examine multiple points throughout the performance process as broken out in the model of individual performance motivation.

1.2.3 Research Question

The research question is the heart and soul of any research endeavor. The question directs the design of the research by supplying a context for methodology decisions. The foundation for the research question is based in the conceptual

framework and constructs of particular interest to the researcher. It is important to make sure that the research question is of enough interest to the researcher to maintain interest throughout the study. The research question for this study is a synthesis of the previously discussed conceptual constructs and the researcher's personal interest in assessment and education.

The following question was posed for this research:

- Can the use of a broad-based holistic assessment method positively affect student motivation?

It is the researcher's contention that using a broad based holistic assessment technique can contribute to improvement in overall satisfaction through the motivation factors of performance.

1.3 Research Expectations

This research produced both tangible and intangible results, which could have significant impact. First is the immediate outputs from the research, or the answers to the questions that are proposed at the outset of the study. Perhaps more important than the immediate outputs are the long term possibilities that are introduced through this research.

It is the researcher's hope that the results from this study will encourage more educators to consider using a more holistic or balanced approach to assessment. The use of purely objective or subjective measurements which have

been used in education during recent years lead to a mismatch between a student's assessment and actual performance.

Another benefit of this research is the application of the balanced scorecard in a unique environment. The more we understand about the interaction between people and performance measurement, the more adept we will become at creating and achieving performance goals.

1.3 Thesis Framework

Chapter 1 (Introduction) introduced the importance of motivation in educational assessment. The chapter also covers the scope of the research, the research purpose and questions, as well as the research expectations. Chapter 2 (Review of the Body of Knowledge) includes a thorough review of the literature surrounding business and educational assessment, satisfaction and motivation in terms of assessment, and the balanced scorecard. The literature review provides a theoretical argument for the use of a broad-based holistic assessment technique in education. Chapter 3 (Research Methodology) describes in detail the procedures and instruments that were used to perform the collection and analysis of data. Chapter 4 (Results) contains the data collected using multiple data sources. The results from the multiple data sources are analyzed. Chapter 5 (Conclusion) is the culmination of the study, where research questions are answered and the data is interpreted. Possibilities and implications for future research are also discussed in this chapter. The Appendix includes examples of the research instruments.

Chapter 2: Body of Knowledge

2.1 Teaching and Grading

Popham defined educational assessment as a formal attempt to determine students' status with respect to educational variables of interest (1995). In the past century, there has been a shift in emphasis from learning and competency to selection and accountability (Reineke, 1998). This shift has far reaching impact into the assessment practices of teachers. The purpose of the assessment has changed from that of guidance to that of judgement.

According to Jenkins, most standard testing practices have four basic flaws (1997). First, student cramming has an effect on weekly reviews and testing scores. Students who excel at cramming can return high marks, making it appear as if they know more than they actually do. Second, exam results often come too late for a teacher to incorporate feedback into their teaching strategy. Tests often mark a conclusion or endpoint of a subject or concept. Third, testing doesn't provide continuous feedback between teachers and students. The test if used alone, is a measure of knowledge at a finite point in time, provides the only feedback mechanism and is unable to contribute process information about student knowledge and improvement. Finally, the most significant flaw with testing is that

teachers are not responsible for the students' learning. The teacher relays information. The students either learn the information, or they do not learn the information. Then the teacher evaluates. The teacher can not make a student learn because they are not in control. In all, there is a failure to address motivation and performance dynamics using typical testing techniques.

So why do students fail? Students fail if they do not satisfy the teacher, or show themselves to be insufficiently prepared to meet the expectations placed on them (Henry, 1992). If a student fails, the question becomes whether the student actually did not know a solution, or whether they simply could not perform on a given testing device. An assessment device is only valuable if it can measure what a student actually knows (Henry, 1992). Behaviors such as cramming can greatly effect the capability of a testing device's ability to measure what a student actually knows.

In the past few decades, there has been increasing debate over assessment practices. There are those that hold that the system as it stands with standard testing forms is the most efficient and reasonable method to assess students. Yet others believe that assessment is a means of persuasion, coercion, and social control by a select group (Rowntree, 1977). Few question the need to know how students are doing. It is simply the method that is the source of debate.

2.2 The Control and Measurement Process

Every day people are asked to make decisions, to weigh options and alternatives, whether it is where to go for lunch, or what kind of car to buy. In a technical sense, the alternatives are measured against a set of criteria and a judgement is made.

In a performance driven world, that is vastly more complicated, it has become necessary to measure performance in order to manage it (Kaplan & Norton, 1996). Getting the right information, to the right people, at the right time has become the art of measuring performance. The purpose of having effective information flow is to allow for effective control of resources. The Plan-Do-Study-Act (PDSA) process can be seen in Figure 2.2.

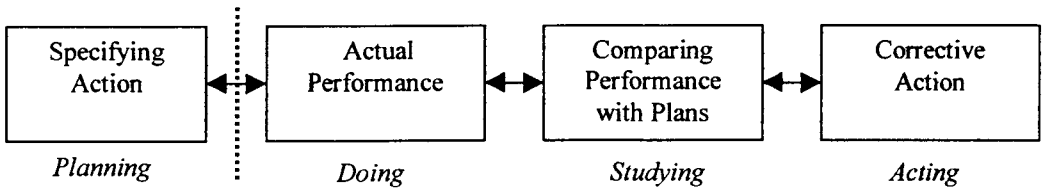


Figure 2.2: The PDSA Process (adapted from Babcock, 1991)

This same process is used in a plethora of functional environments. Of particular interest in this research are business and education. The ultimate link between these two environments are people. Businesses are supplied a continuing stream of new employees that are prepared by the educational system. With such a

tightly coupled relationship between these two groups, you would expect to see a close interconnect between their measurement systems. Oddly, this is not the case.

2.3 Performance Assessment

Performance assessment is a tool applied to achieve goals, to navigate direction, and to create worthwhile and tangible results. Regardless of the environment, there are five dimensions that guide measurement: why to assess, what to assess, how to assess, how to interpret, and how to respond (Rowntree, 1977). As an organization determines what information is needed and when, how it is going to get it, and what is it going to do with it, it must be able to communicate effectively between its individual parts (Kaplan & Norton, 1996a).

An important distinction must be made in measurement systems is the difference between an evaluation and a diagnostic assessment. An evaluation is a status check and looks at only a small performance horizon. Diagnostic assessment looks for patterns over time (Rowntree, 1977).

2.4 The Purpose of Assessing

In both business and the classroom, performance assessments have similar overall goals and purposes. These include selection, maintaining standards, motivating, providing feedback, and preparing for life. (Rowntree, 1995).

The first purpose is selection by assessment and is often considered the strongest link between business and education because often businesses base their employment decisions upon candidate performance in the educational system (Rowntree, 1977).

The second purpose is to maintain standards. There is a high level of performance expectation placed upon people in both education and business. In order to ensure that people perform up to the expected level, performance assessments are put in place to ensure that those expectations are met (Rowntree, 1977).

The third purpose is to motivate the person being measured. People perform to meet expectation, either to gain a particular reward or to avoid a certain consequence. For instance, a student may choose either to complete an assignment to avoid a bad grade in a class or to achieve a high grade (Rowntree, 1977). This is reflected by the sources of power that are designated to teachers or managers: the power to punish and the power to reward (Babcock, 1997). DeVries, Morrison, Shullman, and Gerlach (1981) believe that in business, the two most common uses for performance assessment are salary administration and promotion decisions.

The fourth purpose is to provide feedback to all the parties involved. This is emphasized in the control process as a recurring feedback loop. There are two primary forms of control, open-loop and closed-loop (Babcock 1997). In an open loop system, an approach is set and it is up to the process to determine the final outcome. In a closed loop system, regular feedback is incorporated to continually

adjust the direction being taken. In an educational environment, assessment provides a vehicle that gives both students and teachers feedback (Rowntree, 1997). A collection of surveys cited in DeVries et al. (1981) reported that 57-82% of organizations use their performance assessments to provide feedback to employees as part of their counseling activities.

The fifth purpose is preparation for life, which is more unique to educational assessment. The suggestion here is that performance in the 'real world' is mirrored by the assessment in education and thus it prepares the student to live in an assessment based world (Rowntree, 1977). The parallel between business and education is the identification of development opportunities. Assessment brings out individual strengths and weaknesses.

These purposes constitute the basis for educational assessment but are also the primary reasons to have measurement in a business environment. With the parallels drawn in the discussion above, one would expect to see similar measurement techniques used in both settings, but simple examination of assessment methods clearly indicates that this is not the case.

2.5 Human Factors in Assessment

The psychological effects as a result of assessment are some of the most highly debated issues in the performance assessment literature. There are numerous factors that influence performance assessment including knowledge of the assessment, related extrinsic rewards, competition, evaluation methods, and

reporting methods (Murphy & Cleveland, 1995). This list gives but a sampling of the factors that can impact a performance appraisal and are of particular interest in this study. In educational assessment, students are subjected to a rigorous and seemingly rigid grading structure. A student's particular behavior is governed by, but not limited to, the same factors seen in the performance assessment literature. These include being measured, motivation, competition and ranking, and reporting results.

2.5.1 Being Measured

Students are intimately tied into and aware of the grading structure in their classes. Their knowledge of the assessment can take several paths. It can lead students to align themselves with expectations, or to rebel (Rowntree, 1977). In the first case, the effect is inconsequential. In the second case though, the behaviors assumed may lead to a misleading or invalid assessment.

2.5.2 Motivation

The second factor is extrinsic rewards related to performance. Most reward systems are based on external motivation (Frey, 1997). Business has shifted its focus from hiring employees who do what they are told to recruiting individuals who creatively think about the future needs of the organization (Argyris, 1994).

Historically, managers have relied on extrinsic motivation to get the desired level of performance from their employees. Extensive use of extrinsic motivators leads to crowding out of intrinsic work motivation (Frey, 1997). The more emphasis that is placed on an extrinsic reward associated with a task, the less interest an individual will have in the task (Kohn, 1993a).

Much of the subject of the debate about how to reward centers on the type of motivation used to drive performance. In education, grades serve as the primary basis for the reward system. Herein lies one of the most extreme arguments in performance assessment. Grades and grading in general can reduce the intrinsic joy and satisfaction in learning (Deming, 1994). Using grading as an extrinsic reward is manipulative by nature, and not much different than threatening a particular punishment (Kohn, 1993b).

2.5.3 Competition and Ranking

Competition and ranking are driving factors for assessment especially in education. Students compare themselves with their peers to assess their own performance, and educators compare their students against national standards to assess their teaching methods (Reineke, 1998). Due to the shift from learning and competency to selection and accountability in education, competition for high marks has increased tensions characterized by a large demand for a limited supply of top rankings.

2.5.4 Reporting Results

Finally, the evaluation method and reporting techniques can greatly influence the results obtained from performance assessment. The roots of the performance assessment research are entrenched in measurement and the endless search for new and better techniques to perform assessments (Murphy & Cleveland, 1995). One of the struggles within performance assessment according to DeVries et al. (1991) is alignment and determining the role it plays with the rest of the organization. This is, and will continue to be, a source for the multitude of assessment techniques that are available to the general public. Linking individual performance assessment to organizational goals becomes the next challenge. When individual objectives and goals align with that of the organization, the entire organization is mobilized (Kaplan & Norton, 1996).

2.6 Motivation and Assessment

There is evidence that suggests that an individual's reaction to an appraisal or assessment may be equally as important to long term effectiveness of an assessment system as the validity and reliability of the system's measurements (Mount, 1983). Individuals work within a business system where there are multiple external forces acting upon them. Often, they are not in control and often are controlled by that system. These forces pull an individual in multiple directions resulting in the need to prioritize tasks, expectations, and outcomes. An

individual's motivation determines what outcomes they will strive for and what behaviors they will exhibit in order to obtain those outcomes (Lawler, 1994). Therefore, it is important to consider an individual's motivation within the context of performance and assessment. An integrated approach to individual motivation to work and managing motivational dynamics is presented in Figure 1.2.

The diagram represents the organization and performance components and their relationship with motivation and satisfaction. This model demonstrates a systemic or dynamic view of motivation and performance where feedback from outcomes become inputs into future motivation, thus creating an escalation loop that can be positive or negative (Petrick & Furr, 1995).

The important variables in the model are motivation, individual ability, organizational support, the value and equity of rewards, and end satisfaction. Motivation or a person's willingness or desire to work will be the primary determining factor in the effort that is put forth on a task (Petrick & Furr, 1995). Individual ability is based on the characteristics of an individual and determines if a person is capable of completing a task. Organizational support is the opportunity provided to an individual to perform. Finally, satisfaction is a product of past performance and perceived equity of the rewards. The rewards are a motivating mechanism that are tied with the organizational support or managing system. A disconnect between performance and rewards can be highly demoralizing to the individual (Deming, 1994).

Satisfaction is a function of both the person and the environment (Lawler, 1994). At any one time, satisfaction will be based on the individual's current working situation in addition to their past experiences and expectations of the future. Management systems that incorporate participation and feedback generally result in higher levels of satisfaction and work effectiveness (Petrick & Furr, 1995).

2.6.1 Participation

In general, the goal of an effective control system would be to sustain a positive escalation in motivation. One way to improve performance motivation is to have more participation in work decisions (Lawler, 1986). Participation is most effective at improving motivation when it involves key aspects of the task being performed, such as setting goals (Lawler, 1986). This gives the individual a sense of control and ownership, which in turn improves overall satisfaction according to DeVries et al. (1981). It also increases buy-in for the measurements, and can lead to increased motivation (Spitzer, 1995). High participation in the goal setting and decision processes have also been shown to improve equity (Smither, 1998). By participating, individuals have a greater understanding of the link between performance and the associated rewards.

2.6.2 Feedback

Another key tool that is used to alter performance is feedback. DeVries et al. (1981) state that feedback has been shown to positively influence learning, motivation, and job performance. Feedback is the backbone of the educational process (Rowntree, 1977). Knowledge of past results, goals, and objectives direct and help align individuals with expectations or standards. Feedback reduces uncertainty and provides information for self-assessment (Murphy & Cleveland, 1995). Assessments provide dynamic feedback and form a moment of truth between the superior and subordinate (Lynch & Cross, 1993). The encounter can cause favorable or unfavorable impressions. Great care must be taken to ensure that both subordinate and superior feel that the assessment is valid and fair. Future impressions will be molded by previous impressions (Wagner, 1992).

There are several factors that determine feedback effectiveness. The most important factor is probably the feedback polarity. People crave positive feedback (Spitzer, 1995). Positive feedback keeps employees working towards goals and increases motivation. Negative or corrective feedback can not be ignored, but needs to be framed in a constructive light in order to avoid individual defensive routines. Negative feedback is often taken as criticism, and should be moderated with positive feedback to avoid demoralizing the individual. A series of interviews were conducted with 60 facilities concerning best business practices. According to Longenecker, Stansfield, and Dwyer (1997) those interviewed stated that effective

measurement efforts provide balanced feedback and are not completely critical or negative.

The content of feedback also affects overall feedback effectiveness. Feedback can take on one of two basic content forms: qualitative and quantitative (Smither, 1998). Feedback may simply convey a score that was received, or it may convey behavioral information (Reineke, 1998). It is difficult to capture qualitative information in an assessment, because the accuracy and validity of the data can come into question (Smither, 1998).

Finally, the issue of timing also impacts feedback effectiveness. If evaluation occurs too frequently, then the focus of the job may become to perform well on assessments (Kohn, 1993b). Overall, the frequency and timing of feedback has been the subject of little research. In their study Longenecker et al. (1997) state that 60% of those interviewed said ongoing measurement and feedback systems were used to increase worker productivity. Communication that is frequent and prompt has a motivating effect. A time lag in communication can undermine trust (Spitzer, 1995).

2.7 The Balanced Scorecard

The balanced scorecard is a set of measures that gives top managers a fast but comprehensive view of their business (Kaplan & Norton, 1992). It is a tool that has gained significant attention in the world of business in this decade. The balanced scorecard, provides a method by which goals can be translated into key

operating measures, thereby creating a strategic measurement and management system (Kaplan & Norton, 1996). There are many characteristics of effective measurement systems incorporated into the balanced scorecard.

2.7.1 Four Organizational Perspectives

The balanced scorecard framework includes four primary perspectives: Financial, Internal Processes, Innovation and Learning, and Customer (Kaplan & Norton, 1992). These perspectives can be characterized by asking some of the following questions (Kaplan & Norton, 1992; Skyrme, 1998):

- Financial Perspective- Bottom line: how are we doing? How do we look to the stakeholders? How do we create value for our stakeholders?
- Internal Process Perspective - At what processes must we excel to create superior value for our customers? What drives our organization?
- Innovation and Learning Perspective - Can we continue to learn and to create value? What must we do to succeed in the future?
- Customer Perspective - How do customers see us? What are our customers' needs and how well do we meet them? Are we creating value for our customers?

This framework provides a broad based or holistic view of performance. There are multiple factors that must be considered when you begin the process of combining an organizational strategy and these four perspectives to develop a

strategic measurement and management system. In general, a full deployment of the balanced scorecard has the following characteristics (Silk, 1998):

1. **The balanced scorecard must provide linkage from the vision to strategic objectives to key performance measures – and show 'cause and effect'.** More than a list of measures, the balanced scorecard links organizational strategy and operations through measurable visible mechanisms.

2. **The balanced scorecard must allow creation and linkage of organization and personal scorecards.** The balanced scorecard provides a connection between organizational and personal performance objectives and measurement for mutual progress towards goals.

3. **The balanced scorecard must support both quantitative and qualitative information.** The numbers are important, but the commentaries add real meaning.

4. **The balanced scorecard must encourage dynamic communication.** More than a reporting vehicle, the balanced scorecard is a strategic feedback system. It must support feedback loops, dialogue, comments, personalized assessments and initiative management.

5. **The balanced scorecard must be easy to set up and maintain.** Simplicity is the key to performance measurement. The balanced scorecard has to be able to link goals and objectives clearly, and progress must be easy to track.

6. **The balanced scorecard must be enterprise deployable.** In order for the balanced scorecard to become truly effective, it must be implemented and used at all levels within an organization. This is the only way to create a dynamic feedback between organizational levels.

Kaplan and Norton (1996a) offer a general implementation plan for organizations interested in developing and installing a balanced scorecard. Their steps are summarized in the following list.

1. **Clarify the Vision.** Translate the vision into a strategy that is easily understood and communicated.
2. **Communication and Deployment.** Bring in the next levels within the organization to begin multi-level scorecard development.
3. **Commitment Building.** Build excitement at lower levels of the organization to build buy in and prepare for launch.
4. **Review the Scorecards.** Review of multi-level scorecards for alignment with strategy.
5. **Refine the Vision.** Provide learning and feedback to top level managers who may then need to refine the vision.
6. **Organizational Deployment.** Communicate scorecard to the entire company, and create individual scorecards that link manager's objectives and goals to rewards.
7. **Update Organizational Plans.** Resource planning and allocation issues are reviewed and modified as necessary.

8. **Conduct Regular Reviews.** Review process ensures that what is being measured is contributing to overall strategy, and gives feedback to individuals or groups on performance.
9. **Conduct Top Level Review.** Review process that incorporates lower level measures of performance to help determine overall progress towards vision.
10. **Full Organizational Deployment.** Translation of the organization's performance objectives and goals into a balanced scorecard, rewards and incentives are linked to scorecard.

2.7.2 The Balanced Scorecard as a Strategic Management System

The balanced scorecard now serves as a vehicle for strategic measurement. Businesses combine this measurement system with a form of the PDSA cycle to convert it into a strategic management system. The PDSA cycle can be seen in Figure 2.2. There are four processes in the proposed managing strategy: translating the vision, communication and linking, business planning, and feedback and learning (Kaplan & Norton, 1996c). This process is shown in Figure 2.7.1.

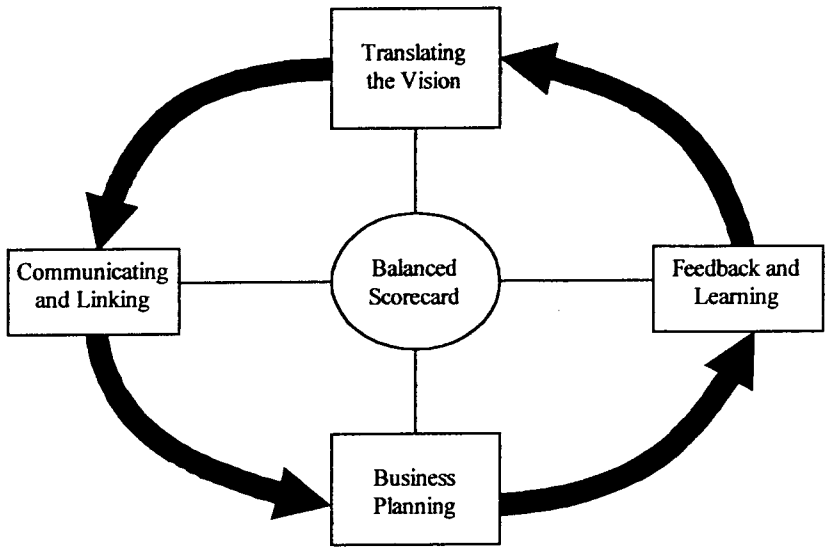


Figure 2.7.1: Managing Strategy: Four Processes (Kaplan & Norton, 1996c)

Phase One is translating the vision and includes the vision being converted into a strategy. This requires that the top-level managers come to consensus on the key goals and purposes of the organization, as well as the key operational measures that define success (Kaplan & Norton, 1996c). The vision must be revisited in the ongoing management process to determine alignment of measures with direction, but also to refine the vision. This also provides an opportunity for a mid-course adjustment of the organizational direction.

Phase Two is communicating and linking which is the process by which the balanced scorecard is deployed across an organization. Deployment is a crucial step in the success of any kind of organizational change. Deployment is the process of participation and feedback that builds a sense of ownership and commitment to the change initiative (Petrick & Furr, 1995).

Phase Three is business planning and is the process through which the strategic management system and resource allocation decisions are linked. This establishes a link between the financial decisions that must be made to keep an organization operational and the strategic vision of that organization. A clear link allows top level managers to drive their organization toward their vision.

Phase Four is feedback and learning, and is the process in which lessons learned and progress made are incorporated into future actions. The balanced scorecard management process emphasizes double loop learning. Double-loop learning occurs when a question or action is taken and the results of the action and reasons behind the action are examined for validity (Argyris, 1994). Double loop learning is a process of examination and reexamination.

The implementation and management processes are linked from the very first step. By adopting an ongoing process of planning, learning, refining, and communicating, an organization can facilitate the implementation process. Each step of the implementation process is designed to follow the management process. This link between the management process and implementation process is represented in Figure 2.7.2. This process utilizes a commitment-based approach to management of the work force (Walton, 1985).

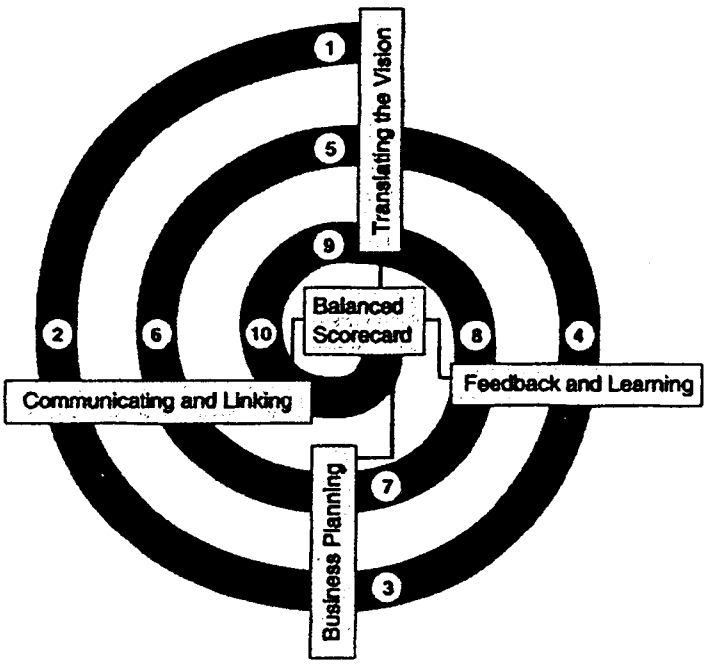


Figure 2.7.2: Using the Balanced Scorecard as a Strategic Management System (Kaplan & Norton, 1996c).

2.7.3 Personal Scorecards

According to DeVries et al. (1981) performance appraisal is the process by which an organization measures and evaluates an individual's behavior and accomplishments over a finite period of time. There is a vast array of literature that discusses an equally vast and expanding number of performance assessment systems. There are several common themes that must be incorporated into an effective performance assessment system.

1. Strategy is translated into clearly defined objectives and goals. According to Chow, Haddad, and Williamson (1997) There has to be a visible link

between an individual's objectives and the overall objectives of the organization.

2. The measures must include a combination of result-based and behavior-based measures. Human Performance has two primary aspects: results and behaviors. Results are the tangible products or accomplishments. Behaviors are the methods and processes employed to achieve results. For all practical purposes, every aspect of performance can be assessed with a combination of results and behavioral measures (Spitzer, 1995).
3. There must be participation on the part of the assessee in the selection of performance measures. Research shows that having input from individuals who will ultimately be assessed by the system during the design phase effects the perceived fairness of and overall satisfaction with the assessment (Smither, 1998).
4. An ongoing process of feedback should be maintained. Everyone associated with the system should know how and when feedback is given. The feedback should have a balance between positive and negative aspects. Feedback must flow both ways. The manager and the employee should take the opportunity to tell each other what has gone well and what has not (Spitzer, 1995).
5. The reward system should be linked to performance goals and objectives. The incentives or rewards must be given in a timely manner, and in an equitable fashion. A focus on short-term objectives as the basis for rewards

will prevent planning and resource allocation to achieve long-term goals (Kaplan & Norton, 1996). The emphasis on rewards should be carefully monitored to avoid crowding out intrinsic motivation (Frey, 1997).

Not every personal scorecard in an organization must be the same. Chow et al. (1997) state that every person brings a different skill set to the table, and therefore their objectives, in light of the overall strategy, may need to be modified.

2.8 Linking the Scorecard to Education

There is a clearly established link between performance assessment in education and in business. In both worlds, it is important that the dynamic process through which people are motivated to work is considered in the assessment design. One of the key variables in the motivation and performance model (Figure 1.2) is organizational support. The assessment device utilized to evaluate performance is a major determining component of other variables down stream in the performance cycle. Issues of support, participation, and feedback are directly related to the systems used to evaluate performance.

The balanced scorecard is a dynamic assessment system that utilizes broad-based holistic assessment methods. It utilizes a high level of feedback over a wider range of process and performance measurements. This is part of the reason it has seen such wide success in business. This is also why it has so much potential for other fields like education. Most educational assessment techniques have multiple flaws when it comes to assessing student knowledge. It is very much like trying to

evaluate the performance of a person who paints houses. You can not watch every stroke because it is too intensive. The use of single point-in-time tests is similar to evaluating performance on doors and windows only. You can't evaluate based on the number of houses completed because the quality of the paint job could be poor. What gets measured is what gets done. Therefore, how do we evaluate the performance of the painter? How do we measure the performance of the student?

Applying a balanced scorecard to educational assessment has the potential to improve student performance assessments. The success of the balanced scorecard in many environments suggests that the broad-based holistic style of measurement should have higher probability of capturing a student's true performance. The feedback mechanisms inherent with the balanced scorecard approach will provide both students and instructors with more information about their strengths and weaknesses. Finally, the most important objectives and processes can be directly targeted with measures designed to encourage or direct performance. By capturing a more a clear picture of performance, there is a higher chance of equitably rewarding students, thus increase their satisfaction. The following chapter details the methodology developed to test this premise.

Chapter 3: Design Issues and Research Methodology

The purpose of this chapter is to set forth the design issues and research methodologies that were used in this research. This chapter is divided into two parts. The first section details the design issues and decisions required for this research. The second section defines the research methodology that served as a framework to support the research activities.

3.1 Design Issues

Many aspects of motivation can not be captured in a controlled, direct, and quantitative fashion. Therefore, a mixed quantitative and qualitative approach to the research was required to study the phenomenon of interest, which was student motivation. Table 3.1 presents the design issues, alternatives, and decisions for this research application as adapted from Van Aken (1991). The following sections provide detail about each design issue and the rational for how and why each design issue was resolved.

Table 3.1: Design Issues and Options (adapted from Aken, 1991)

Issues	Sample Options and Considerations	Option/Decision for this Research	Where Discussed
What is the primary purpose of this study?	Basic research, applied research, summative evaluation, formative evaluation, action research	Action research Summative evaluation, formative evaluation,	Section 3.1.1
What is the focus of the study?	Breadth versus depth	Low breadth, high depth	Section 3.1.2
What are the units of analysis?	Individuals, groups, program components, whole programs, organizations, communities, critical incidents, time periods, and so on	Individuals enrolled in two college courses.	Section 3.1.3
What will be the sampling strategy?	Purposeful sampling, probability sampling, variations in sample size and makeup	Combination strategy of purposeful sampling	Section 3.1.4
What types of data will be collected?	Quantitative, Qualitative, or both	Both	Section 3.1.5
What controls will be exercised?	Naturalistic inquiry, quasi-experimentation, experimental method, case studies	Quasi-experimentation, naturalistic inquiry	Section 3.1.6
What analytical approach or approaches will be used?	Inductive versus deductive analysis	Deductive analysis	Section 3.1.7
How will validity and confidence in the findings be addressed?	Construct validity, internal validity, external validity, reliability, data triangulation	Construct validity, data triangulation	Section 3.1.8
When did the study occur and how was it sequenced?	Long-term fieldwork, rapid reconnaissance, exploratory phase to confirmatory phase, fixed times versus open timelines	Rapid reconnaissance, exploratory phase, fixed time	Section 3.1.9
How were logistics handled?	Gaining entry to the setting, access to people and records, contracts, training, endurance, and so on.	Personal contacts, voluntary support	Section 3.1.10
How were ethical issues and matters of confidentiality handled?	Informed consent, protection of human subjects, reactivity, presentation of self, and so on	Informed consent, university research approval	Section 3.1.11
What resources were available?	Personnel, supplies, data collection, materials, analysis time and costs, reporting/publishing costs	Expenses covered by primary researcher	Section 3.1.12

3.1.1 What is the primary purpose of the study?

This research addressed the issue of student motivation in education as it relates to assessment. This research has characteristics of several of the different research types. Its origin was in action research. Traditional grading and assessment mechanisms have resulted in the gradual decay of student motivation or joy of learning (Deming, 1994). Education has turned into an obstacle to overcome in order to open a door to a better future, rather than an opportunity to better oneself and become a more productive and contributing member of society. The intention of this research was to find a way to reverse this decline in student motivation through assessment methods.

This research also has aspects of formative evaluation. The underlying social structure and student behaviors in education are not well understood, and are critical to student motivation. This study sought to understand and answer some of the questions about student motivation that affect the classroom learning environment. The information collected could be used to aid the design of future performance assessment methods.

Finally, this research also included aspects of summative evaluation. To help support the collection of useful information, an intervention will be applied. This intervention was in the form of a broad-based holistic performance measurement system called a balanced scorecard, which was described in Section 2.7. The intent of this intervention was to provide a condition to study the effects of assessment on student satisfaction and motivation.

These three forms characterize the nature of this research. The research does not directly match the conventional research types because the phenomenon of interest is very difficult to isolate from other issues related to educational assessment. Therefore, a combination of the different types was adapted to better address the research focus.

3.1.2 What is the focus of the study?

The challenge of conducting research is determining the boundaries of the phenomenon that will be studied. The initial topic is typically too large and insurmountable, and must be refined and remolded into a feasible study. As with any kind of research, there is a trade-off between the depth and breadth of the study that is a result of trying to distribute limited resources.

This research deals with student attitudes and motivation, which contains an extremely vast array of research possibilities. It was important to be able to focus the study in order to develop substantial understanding of a particular phenomenon. Therefore, in this study the breadth or number of phenomenon being studied within general motivation was reduced and the depth of the phenomenon of interest, motivation and satisfaction, was increased.

3.1.3 What are the units of analysis?

The units of analysis for this research were the students enrolled in two senior/graduate level college courses. These courses were scheduled courses at Oregon State University during the winter term of 1999. The selection of the study courses involved multiple criteria. One course served as a control, and the other as the experimental case. In order to draw conclusions, there needed to be significant similarity between the courses. This symmetry allowed comparison of measured results between the courses. Issues such as class structure, professor, student enrollment, student characteristics, course topic, testing, and assignment load were taken into consideration to ensure comparability.

3.1.4 What will be the sampling strategy or strategies?

As is typical with qualitative research, a purposeful sampling strategy was deployed in order to obtain find information-rich cases to study. In a university setting, there are a large number of cases to choose from, so determining which would fulfill the needs of the research is very important. A combination or mixed strategy of purposeful sampling was deployed to determine which courses would be used.

Operational construct sampling was the primary strategy employed when choosing the cases. It was determined that the best way to study student motivation would be to use real world course situations. Using university courses allowed the

researcher to track motivation throughout the duration of the term in an environment that was familiar to the study subjects.

Opportunistic sampling was another strategy used. As discussed earlier, course symmetry was very important in this research. This study required a high level of cooperation from instructors and students in order to obtain data. Voluntary participation on the part of the instructors of courses was utilized providing more flexibility.

Convenience sampling was also employed. The time available was a limiting factor in performing this study. Therefore, it was important to take advantage of instructors with whom the researcher was already familiar, and who would be readily available for consultation.

3.1.5 What types of data will be collected?

A mixed form of data collection was employed in this research. This study utilized a quasi-experimental design, qualitative data, and content analysis. In the research, one course served as a control, and the other as the experimental course. The purpose was to be able to collect information in order to compare the experimental course with the control course.

Within the courses, multiple sources of data were collected. This allowed for data triangulation within the research. In land surveying, triangulation is the use of multiple points in order to determine where a single point or location is. Similarly, by using multiple data sources, the results have a stronger foundation of

data to support conclusions. Data for this research was collected by a series of surveys, direct observations, and concluded with a series of focus groups with the subjects in the experimental course. Detailed descriptions of the design and application of these data collection methods are discussed in section 3.2.3.

3.1.6 What controls will be exercised?

From the outset of research, it was important to determine the functional approach to the research methods. On the extreme ends of the spectrum are naturalistic inquiry and experimental evaluation. In naturalistic inquiry, the research exerts no controls over the experimental environment, thus allowing events to naturally unfold. In experimental evaluation, the research setting is completely controlled allowing for no variance between subjects.

This study fell in the middle of these two extremes. In pure naturalistic inquiry, no form of intervention would be used because this causes a shift from the natural environment. In pure experimental evaluation, nothing is left to chance, or to develop naturally. Therefore, a quasi-experimental form of research was applied in this study. This gave the opportunity to begin the study in a controlled experimental condition, and then allowed events to unfold naturally over the duration of the experiment.

The controls exerted in this study occurred primarily in the selection of the study courses. Selection of the courses based on similarity required selection of criteria resulting in as controlled a condition at the outset of the study as possible.

Data collection procedures were also controlled and unchanged throughout the duration of the study. Beyond this, all other events were allowed to unfold naturally as the course progressed.

3.1.7 What analytical approach or approaches will be used?

In research, there are two methods of analysis by which researchers attempt to come to conclusions: inductive mode and deductive mode. In a deductive mode of analysis, researchers attempt to test a theory in order to come to a conclusion. In an inductive mode of analysis, researchers try to come to conclusions by observing activities and identifying themes without any preconceived expectations.

This research utilized a deductive mode of analysis. A strong theoretical foundation was developed from the motivation and measurement literature. Therefore there was an expected result from the outset of the experiment. Hence, the research data collection was designed to target the specific phenomenon of interest.

The analysis method used a combination of statistical and content analysis. Comparison analysis between the experimental and control courses was performed by testing for statistical significance between the study groups. In addition, the use of qualitative data collected through focus groups was used to explain unique events and clarify attitudes.

3.1.8 How will validity and confidence in the findings be addressed?

Research is only valuable if the methods used to collect data and come to conclusions yield valid results. In every research endeavor, the challenge becomes developing robust and defendable methods to assess a desired phenomenon. Researchers attempt to establish four types of validity: construct validity, internal validity, external validity, and reliability.

Construct validity is described as the ability to establish measures that correctly measure a concept or phenomenon of interest. Internal validity addresses how accurately the information collected is, and how well it fits the phenomenon being studied. External validity establishes the domain to which a study's results can be generalized too. Reliability is a level of confidence in the ability of the results to be duplicated or generalized to another context. These issues are addressed by using multiple data sources and triangulation. Depending upon a single source of data can be problematic. If the measurement tool is inappropriately constructed, then this can seriously effect the validity of the results. By using multiple sources of data, you approach a phenomenon from multiple directions. If multiple sources of data converge to a similar conclusion, then this suggests that a particular phenomenon has been captured appropriately.

This study utilized a survey as a measurement instrument. Surveys are useful instruments, but they are only useful in a research endeavor if they accurately measure useful data. In this study, a survey was adapted from the Morse Indexes [sic] of Employee Satisfaction. Validity and reliability information for the

Morse Indexes of Employee Satisfaction were available and showed a good level (r=.35 to r=.52) of intercorrelation, suggesting high internal consistency (Miller, 1991). The survey employed in this research would be expected to have similar characteristics. Utilizing a validated survey instrument as a base should improve the reliability and validity of an adapted instrument.

Direct observation and focus groups were also used to collect qualitative data to support findings. The direct observation was used to collect data on participation. This measure targets willingness to work or perform. Focus group data was collected at the end of the course using three groups. The data was internally triangulated to ensure validity, and then triangulated with the other data sources.

External validity is the ability to generalize findings to a general population. To a certain extent this is possible in this research. Many of the characteristics of the study sample can be carried across to all university college courses. However, there are some limitations to the generalizability of the findings. The primary limitation is course content. For the purpose of illustration, consider the differences between a creative writing and a calculus course. The level of subjective and objective assessment standards varies greatly in these two types of courses. In a university setting, there is a plethora of different courses, each with it own unique assessment requirements. While it is possible to compare the sample course to many courses in the university, it would be difficult to predict the outcome with every course.

3.1.9 When did the study occur and how was it sequenced?

Timing and logistics was the most limited aspect of this study. To study the effects of motivation in the context of a course allows only one term (assuming a term based system) or eleven weeks for study. In this time, students theoretically go from no assessment (no performance information) to final assessment (permanent grade). It was the intention of this research to understand how student satisfaction changed over time. The problem is that with such a limited window of opportunity, it is difficult to unobtrusively study the phenomenon.

It was determined that a survey would be issued five times during the course of the term. These would be spaced every two weeks in order to track trends over time. This design led to the use of two courses because the effect of filling out the surveys would create a bias in the data. By using two study courses, the information could be compared, and the bias effect of collecting data could be filtered out.

Direct observation was utilized to collect participation data. This data was collected for every meeting of both courses. The number of class periods and total time differed for the two courses. The primary researcher attended every class meeting of both courses to collect the participation data.

The focus groups were conducted during the final week of exams after all course meetings were complete. They were conducted over the course of two days.

The primary researcher and two assistants conducted the focus groups to collect qualitative data about the experiences and feelings of subjects during the term.

3.1.10 How were logistics and practicalities handled?

The sample courses were selected by presenting the research idea to two instructors, who then agreed to participate in the research. These two instructors then became focal points through which survey efforts were coordinated. On scheduled days, the primary researcher would bring in and distribute numbered surveys to the study subjects. They would be given a few minutes to complete the survey, and then return it.

The primary researcher attended every course meeting in order to obtain direct observation for the participation data. This step required the researcher to learn to match the names and faces of each study subject so that when they spoke, their participation could be recorded. The courses used nametags during the first week to facilitate this process. A standard template was used to track participation data.

The focus groups were held at the end of the term after all class information was disseminated. Participation in the focus groups was completely voluntary so an incentive was used to encourage participation. The focus groups were held at a local restaurant to provide a more open atmosphere. The primary researcher and two assistants were present at the focus groups. The primary researcher asked a set of predetermined questions and listened carefully to the responses. One assistant

was a dedicated note taker and used a large flip chart to capture responses from the students. Subjects were asked if the written notes corresponded with the meaning of what they said. The second assistant provided service by getting non-alcoholic beverages and food for the participants so that they would not have to leave and disrupt the group. A tape recorder was used at the first two focus groups, but background noise made it almost impossible to create a transcript.

3.1.11 How were ethical issues and confidentiality handled?

All research at Oregon State University that utilizes human subjects must have the research methods approved by the University Research Board. A consent form and a sample survey were submitted to the research board for approval. Approval was granted for a one-year study period.

Subject selection was not radomized. Subject selection was confined to students who enrolled in the two subject courses. Participation in the study by students in the course was not mandatory, and students were unaware of the study until the first course meeting. At the first course meeting, students were told that there was going to be research conducted in the course and the nature of what would be required if they chose to participate. They were not told any background or expected outcomes of the research. Students were told that they did not have to participate if they did not want to, but in both courses, all students in attendance at the first course meeting agreed to participate.

The primary researcher introduced himself at the first course meeting and explained that he would attend all of the class meetings. At no time were the data collection procedures changed or modified. The researcher did not examine any survey data until all data for the study had been collected to prevent researcher bias. At no time was any of the collected data shared with the test subjects. Subjects were not given any information about the purpose or content of the research other than the survey and focus group questions themselves.

The informed consent form was given to each subject accompanied by an explanation of how the research would and would not be used. Students were issued a subject number as part of the consent form. All information collected by survey was identified by subject number only. Any data later shared with the instructors would be on a subject number basis, not by student name.

All data that links the name of the subject with their responses was destroyed at the completion of the study. At no time during this study did any subject come forward to express concern with the data collection, and subjects validated all information taken during the focus groups.

3.1.12 What resources were available?

Time was the most limiting factor of this research. Because of the relatively short duration of the college term, and the other commitments of the primary researcher, it was very difficult to schedule research time.

The instructors in both classes also provided the opportunity to perform this research as a part of their courses. The instructors for the experimental course were extremely valuable. Their willingness to adopt the assessment intervention that was designed and their efforts to implement it were invaluable to this research. Without their voluntary cooperation, this study could not have been possible.

The materials required for this research were minimal. There were a total of 50 subjects in the study over two courses. Total printing costs for surveys and participation templates was less than $100. The university department provided a flip chart at no cost. The focus group participants were treated to pizza and soda at a local restaurant for a cost of $175. The tapes and tape recorder were previously owned by the researcher and were utilized at no cost.

3.2 Application of Design Issues in Research Methodology

The following sections describe the design issues and general framework of the research conducted. This portion of the methodology is intended to describe the designed components of the research that was conducted.

3.2.1 Establishing the Research Idea

Coming up with the initial research idea and then developing the research methodology is a difficult and often daunting task. In this respect, the idea for this research was developed with relative ease. The foundation for this research was

grounded in the performance measurement and motivation literature. These areas were of particular interest to the primary researcher. The researcher drew upon previous industry and academic experience in measurement systems.

The initial link between the balanced scorecard measurement system and educational assessment was established in the context of graduate level cross-disciplinary course. A team composed of two educators, a health care professional, and the primary researcher (an engineer) developed a case that highlighted the need for a change in educational assessment. One of many factors addressed in this effort was student motivation.

3.2.2 Operationalizing the research

The most difficult part of this research was establishing a sample population to study. In a qualitative study, determining the sample and determining the method can be a chicken and egg problem. This is why many qualitative research methods are adaptive and subject to change due to variations in the research environment. In addition, the sample had to fulfill the similarity requirements that would automatically exclude some courses from consideration.

3.2.2.1 The Sample

It was determined that using an academic college course would provide the best opportunity to study the phenomenon of student motivation. The phenomenon

of student assessment is a cyclical process that begins and ends with a university assessment period, in this case an eleven-week term. Therefore, any research effort would have the limited horizon of one assessment period.

The use of a quasi-experimental design, utilizing a control and experimental group required the use of multiple courses. The concept of mixing control and experimental subjects in the same course was disregarded because of the nature of the research. While it would increase the validity of the results, it would be unfair for the subjects of the same course to be evaluated with different assessment procedures.

Selection of multiple courses was also very difficult. To obtain the best experimental design, it would be optimal to use the same course, with the same professor, and the same experimental subjects or completely random subjects. The experimental and control group subjects would be homogeneous, exhibiting similar characteristics and background. The instructor would present identical material, in an identical fashion to both groups. The real world poses several infeasibilities for this type of a sample on the part of the students, instructors, and the researcher.

Students are not required to take the same course twice, thus forcing a renewed population of students. Students also have a purpose in attending a college course and therefore subjects are not randomly selected. Asking student peers to undergo different assessment methods within the same course introduces a series of additional inequity conditions that could seriously effect the results of the study.

Using the same instructor for multiple sections also poses a problem. Due to the size of the university, a single professor rarely teaches two sections of the same course. When it does happen, asking an instructor to use two different assessment techniques creates an unfair burden, and leaves them vulnerable to scrutiny by peers and students. Instructors can not present material identically to two different groups. The exception to this is a TV or distance learning course. In this case, while material is the same, the teaching medium is not, and could result in potentially significant sources of variance.

Finally, most courses are offered only once a year. Therefore, to use the same course would require a year interval between collection of the experimental and control data, which could result in a number of unforeseeable variances. It would also significantly increase the cost of the research

Voluntary participation by course instructors was an important factor in the success of this research. The balanced scorecard provided a foundation for the intervention that would be used in the experimental class. To adapt this performance assessment method to a course environment required a great deal of flexibility and pre-planning on the part of both instructor and researcher. This up-front design work was very important, because after data collection began, the researcher could not continue to contribute to the design without sacrificing data integrity. In addition, the intervention also called for the participation of the students in the design of the assessment method after the primary researcher was no longer participating. Therefore, it was important to find willing instructors who

were both flexible and who could adapt to changes in the initial intervention design.

With these issues considered, it was determined that two separate courses during the same assessment period would be used. Potential courses were examined for voluntary participation and for symmetry of criteria including class structure, instructor similarities, homework assignments, testing, projects, and course content. The goal was to find two classes that were as similar as possible while still being able to attain instructor consent. Two courses were selected with regard to these criteria: IE 470/570 Management Systems Engineering, and IE 491/591 Selected Topics in Design for the Environment, Safety, and Health.

Both courses presented major similarities, and were owned by instructors who had previous contact with the primary researcher and who were receptive to helping in the research effort. Major similarities included class structure, student background, instructor similarities, and assignment layout. The student enrollment was made up of junior, senior, and graduate level students in each class. Both courses divided topics on a per week basis and included a weekly writing assignment based on the topic. Both courses utilized two student projects, compromising a major component of the course work. Both courses utilized a team of two instructors, providing a fair and unbiased assessment in both courses. The only major dissimilarities were that the course subjects and instructors differed. In the case of course subject both were topics based courses. In the case of

instructors, the teams had similar discussion/participation oriented teaching styles. Also, both courses were being offered for the first time.

3.2.2.2 The Intervention

At the heart of this research was the desire to understand the link between assessment methods and student motivation. One of the vital aspects of choosing the sample was the similarity between course work and content. This ensured that, when the intervention was made, a valid comparison between the experimental and control groups could be performed.

Aside from the uncontrolled variables, what set the experimental and control groups apart was the assessment method intervention. In the experimental group, an assessment method called a balanced scorecard was utilized to measure student performance in the course. The balanced scorecard used in the experimental course was an adapted form of the balanced scorecard presented by Kaplan and Norton (1993). To facilitate the process of adaptation, the course instructors in the experimental course combined assessment planning with course planning.

One key feature that sets the balanced scorecard apart from many performance measurement systems is the emphasis on participation and feedback. The instructors set out the primary course goals and objectives. In the design and selection of performance measures, students were asked what they felt were important measurement criteria for the goals and objectives of the course. This key

phase allowed students to have input into the measures they would be evaluated with as well as increased their ownership in the final assessment method. It should be noted that the research measures were independent from the balanced scorecard measures. The focus of the study is to understand how the broad-based balanced scorecard assessment technique compared to the traditional assessment practices in the area of student motivation.

The other key feature that differentiates the balanced scorecard assessment technique apart from traditional assessment practices was the midterm evaluation review. This relates to linking the measurement system to feedback mechanisms. In traditional assessment practices, assignments are returned with some sort of alphanumeric indicator of performance, perhaps with some comments. In the balanced scorecard technique, both instructors met with each student to discuss performance to that point in the term. Before this evaluation, students were asked to fill out a self-assessment scorecard explaining how they felt they had performed. Similarly, the instructors filled out a similar form for how they felt each student had done. At the midterm review, a dialogue provided both instructor and student with feedback for how the course was progressing, identifying strengths and weaknesses, and areas for suggested improvement for the remainder of the term.

Theory suggests that by providing an opportunity for students to participate, and including a high level of feedback on performance, that student motivation should increase. The purpose of the intervention was to develop and demonstrate

an assessment method with high feedback and participation to set a framework for studying student motivation.

3.2.3 Data Collection

The research used multiple sources of data in order to develop a valid case for the study conclusions. These data sources sampled the same group of subjects with three different techniques: survey, direct observation, and focus groups. Each method required careful consideration and planning to determine what would be asked, observed, and examined with each data collection technique. In order to insure validity, each data collection method had to be carefully designed and administered in order to accurately capture the desired constructs.

3.2.3.1 The Survey

The primary measurement tool in this research was a survey. The use of a survey can provide a large amount of information about a topic. The survey is a quantitative method used to obtain information on constructs that are not be physically measurable. Typical quantitative designs contain five components: the quantitative method (survey), population and sample, instrumentation, variables, and data analysis (Creswell, 1994).

The purpose of utilizing a survey was to be able to collect information about subjects' attitudes. Measuring motivation in a quantitative fashion is very difficult.

Motivation is an abstract construct of human attitude and therefore can not physically measured or directly examined. Hence, human motivation is very difficult to assess. The use of a survey is one of very few quantitative methods of assessing human attitude. Population and sample were discussed in sections 3.1.3, 3.1.11, and 3.2.2.1.

As stated previously, the survey used in this study was adapted from the Morse Indexes of Job Satisfaction. The choice to use a previously designed survey was to increase validity and reliability. Morse's survey was designed to be used as an employee interview at a work-site. In using a survey, several design issues had to be considered: adaptability, measured constructs, and applicability.

In Figure 1.2, a model of dynamic motivation and performance was presented containing attitude, performance, and organizational factors. Individual motivation is a cyclical process where the outcome from previous processes becomes a input in the next cycle. Consider a university course where each term begins with a new assessment. Performance, satisfaction, and other aspects from previous courses become attitudes that a student will take into the next course.

The questions of the original Morse Indexes had to be compared and adapted to the components of the motivation and performance model. Table 3.2 maps the components from Figure 1.2 to the survey questions. The heart of this research was student motivation and we can relate individual survey questions back to the motivation components. The original indices included Company Involvement, Financial and Job Status, Intrinsic Satisfaction, and Pride in Group

Involvement (Miller, 1991). The first three indices were adapted to a classroom environment. The final index required a more extensive revision from group to individual work. The resulting survey covers the major variable components in Figure 1.2. The final survey can be viewed in Appendix A.

Table 3.2: Relationship between Research Survey and Motivation Components.

Questions	Motivation Willingness to Work	Reward Equity	Individual Equity	Satisfaction
1. How well have you like this class?	X			
2. Would you advise a friend to take this class?	X			
3. Describe how you feel this class is going? Do you feel you have been a part of the class?	X			
4. How well satisfied are you with your current class standing?		X		
5. How well satisfied are you with your chance of improving your evaluation in this course?		X		
6. How satisfied are you with the way things are going in this class?		X		
7. In this course, have you felt frustrated with your ability to progress towards your or the class's goal? If so, describe an example.		X		
8. How well did you like the sort of class work you were doing?				X
9. Do you feel the assignments give you a chance to do what you do best?				X
10. Do you feel a sense of accomplishment from the work you are doing?				X
11. I feel my contribution in this course (of)_____?				X
12. How well do you think you compare with others in the class at getting things done?			X	
13. How well do you think you compare with others in the class in quality of work?			X	
14. How will do you think you compare with others in contribution to discussion?			X	
15. What level of identification do you feel with the others in class?			X	

The survey was administered five times over the duration of a term. The first survey served as a baseline and asked questions about previous course experiences. The next four administrations focused on the particular course being studied. This layout served to establish an initial comparison between the two courses being studied.

The analysis of the survey data was primarily based on comparisons between the experimental and control group. The researcher expected that there would no difference between the average responses on the baseline survey, and that a significant difference in motivation would be identifiable between the experimental group and control group in subsequent surveys.

3.2.3.2 Direct Observation

One of the ways in which people learn is by participating and observing something. Therefore, it is realistic to expect that one of the best ways to study a phenomenon is by physically observing it. In this research, direct observation of every class meeting was utilized to collect information for this research.

In the course of structuring direct observation, five basic issues must be considered: the role of the observer, portrayal of the observer, portrayal of the purpose of the observation, duration of the evaluation, and focus of the observations (Patton, 1990). The observer was a partial participant. While not

registered in the course, the observer occasionally participated in group discussion. With the direct observation, the research effort was direct and up-front to the study subjects. All subjects were aware that participation and attendance information was being observed for the research in class.

Direct observation data was collected during every class meeting. The observer would begin each class period by taking a roll. This was done non-verbally. During all large group discussions, participation data was collected. Participation data was defined as every time a student made a substantial or insubstantial comment. This required that the researcher learn and match the names and faces of all the study subjects. In addition to the participation and attendance data, the observer also recorded major daily events that might be considered as special cause events in the participation data.

Analysis of direct observation data was done by statistical comparison between the experimental and control groups. The data was compiled into total amounts of participation per week in a quantitative fashion. The qualitative special causes of data variation were only used for data interpretation. Due to the inclusion of participation in the assessment model for the experimental course, it is expected that higher satisfaction will yield higher participation. The cyclical nature of the motivational model suggests that the willingness to work (in this case participate) should increase with a higher level of satisfaction and other components in Figure 1.2.

3.2.3.3 Focus Groups

The final source of data for this research was a series of focus groups. A focus group is a type of non-directive group interviewing technique that is a specifically useful for assessing attitudes and personal feelings of people about a specific topic. In this study, three focus groups were conducted at the end of the course to try to capture student attitudes and reactions with respect to the balanced scorecard intervention in the experimental course.

The focus group sessions addressed eight questions. The process of selecting questions for the group required consideration of several factors. During the course of the term, subjects could have developed very strong feelings towards the intervention. Therefore, the order and questions themselves were designed to allow them to vent their positive and negative feelings in the first few questions in order to prevent strong emotions from biasing their answers. Other questions were associated with measures of satisfaction and equity associated with the balanced scorecard assessment method. The questions in the focus group were purposely open-ended and did not directly target any one motivational component, but allowed the study subjects to explain factors important to them. The focus group instructions and questions can be seen in Appendix F.

Setting up and administering the focus groups was an important step in the focus group process. One of the important steps in the administration of a focus group is to determine the location or setting in which the focus group will be conducted. Since the focus group was held after the regularly scheduled term and

participation wasn't mandatory, it was necessary to provide an incentive. Hence, the focus groups were conducted at a local pizza parlor and free pizza and beverages were provided.

As previously mentioned, participation in the focus groups was voluntary. During the last two course meetings, a sign up sheet was circulated among the students. The students of the course were consulted to determine what times would be appropriate to hold the focus groups. It was determined that they would be held on the evening of Tuesday and Wednesday following the last meeting of the course on Monday.

Within a focus group, protocol and purpose must be laid out. First, the purpose of the focus group must be clearly relayed to the subjects. This was achieved by reading a prepared statement at the outset of the focus group. Subjects were informed that the relationship between the students and the balanced scorecard assessment was the focus of the group discussion. They were told that there were no right or wrong answers, and that all information would be kept confidential. Subjects were also informed that what they said would be recorded in both written and tape formats.

A structured process must be followed in conducting the focus groups. The tape recorder was started in advance of asking the first question. There were three primary roles that were required to conduct the focus groups. These include the questioner/observer, a note taker/observer, and an attendant. The questioner/observer's job was to facilitate the session by asking the questions and

observing the reactions and answers of the subjects. The note taker recorded all responses by the group on a large flip pad. This served as both a group working memory for the subjects in the group and as a way for the researchers to ensure the actual ideas and comments of the subjects were captured. The attendant was a courtesy roll whose responsibility was to run between the group and restaurant personnel to get food and drinks for the subjects. This prohibited the participants' food and beverage needs and desires from being a distraction to the discussion.

Analysis of focus group data can follow four different options (Krueger, 1994). These options include transcript-based analysis, tape-based analysis, note-based analysis, and memory-based analysis. For this research, a note-based analysis was selected for this study. Transcript and tape based analysis are the most thorough forms of analysis in this type of research. They are best suited for research where they are the only form of data being collected, and where they are the only way to disseminate focus group conversations. The note-based analysis relies mainly on field notes taken during the focus group. This study used a group working memory and flip chart approach to note collection instead of typical scribing. This allowed us to verify comments as the focus group was conducted. This is an appropriate method of analysis where multiple data sources are being utilized to reach research conclusions.

The notes collected in the focus group were sorted for content and major themes. Triangulation of themes between multiple focus groups was used in order

to increase confidence and reliability in the conclusions drawn from them. Triangulation also increases confidence in the data validity.

3.2.4 The Final Analysis

The research methodology utilized a combined or mixed form of data sources including a survey, direct observation, and focus group data. For each data source, reliable collection procedures were utilized in order to obtain valid data on the research construct. To increase validity of the results and final conclusions, triangulation between each of the three data sources was used in combination with theoretical constructs. Since each of the data sources targets the relationship between motivation and the assessment, the methodology provides sound data for the theoretical and empirical analysis, which is presented in the following chapter.

Chapter 4: Results and Analysis

4.1 Chapter Overview

This chapter presents the results and analysis of data collected during a 10-week college term. The data was collected using a series of surveys, direct observation, and focus groups. The first section of this chapter presents time-series survey results. This was the most important data component and clearly shows changes in student attitude during the course of the term. The direct observation data shows changes in student participation over the course of a term. Qualitative focus group data is presented and used to clarify reactions to the intervention and to explain sources of special cause variation within the survey data in the experimental group.

4.2 The Survey Analysis

The survey questions were broken down into four indices. These indices map to motivation, reward equity, individual equity, and satisfaction. Table 3.2 shows how the survey questions map to the four performance components in the motivation model presented in Figure 1.2. The survey contained fifteen questions. Thirteen of the fifteen of the questions were based on a five point ordinal scale. The remaining two questions were open-ended and asked for detailed data to

support the trend data in a category. Three questions were related to motivation, one of which was open-ended. Four questions were focused on reward equity, which contained the remaining open-ended question. Four questions dealt with individual equity. The remaining four questions addressed individual equity.

The survey results on quantitative questions are broken out by component in sections 4.2.1 through 4.2.4 and shown graphically as line charts in Figures 4.2.1 through Figure 4.2.4. Presented in the body of this thesis are the grouped construct results. The individual question results are located in Appendix B through Appendix E. Each graph was has two lines, one each for the experimental and control groups. Each line is composed of five points, one for every time the survey was issued during the duration of the term. The first point served as a baseline for the study. It asked about previous course experiences. It is expected that the average response on the survey questions will be the same for both study groups on the first survey due to similarities between groups. Results from qualitative questions on the survey are summarized in Tables 4.2.2 and 4.2.4.

The response between groups was statistically compared for significance. The average response between groups was statistically compared using a t-test for a normal distribution, with variance unknown. The experimental group had 34 test subjects, and the control group contained 13 subjects. F-tests were used to compare variations between the sample variances. Finally, inter-item reliability was assessed using Cronbach's alpha. SPSS Graduate Pack 8.0 for Windows was used

to calculate inter-item reliability for groups of survey items. The following four sections break down response analysis by concept group.

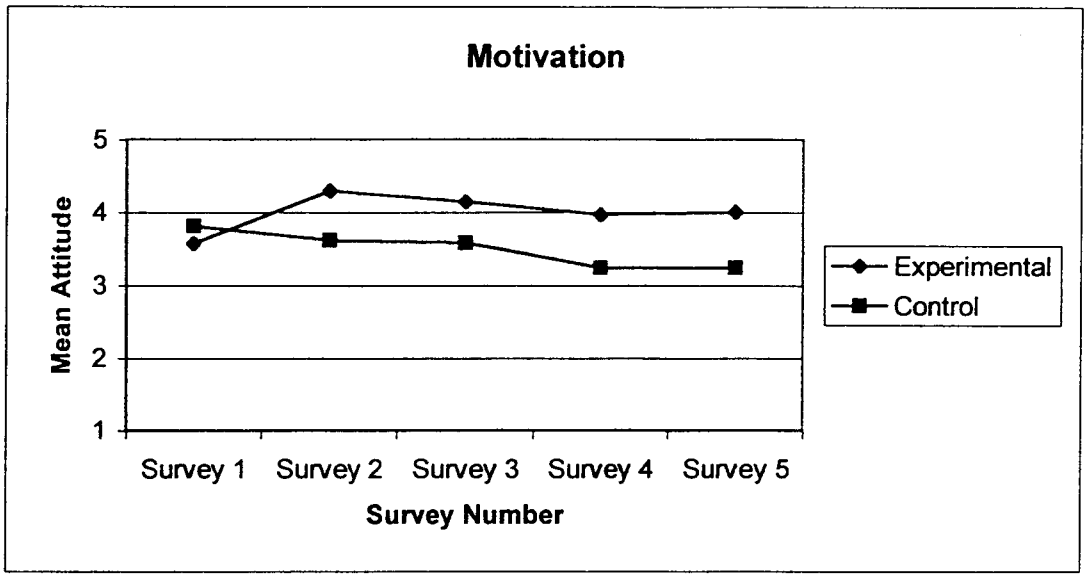
4.2.1 Motivation

Intrinsic satisfaction is derived internally from past experiences and is manifested in the desire or willingness to perform in future situations. Increasing the intrinsic satisfaction of a student can increase the level of motivation with which an individual will approach future tasks. The first group of questions targeted intrinsic satisfaction of the students.

Figure 4.2.1 is a graphical representation of the survey results on the motivation construct. The experimental group's motivation was higher than the control group. The statistical analysis results are presented in Table 4.2.1. Consistent with the proposed theory, the results suggest that student motivation was higher in the experimental group. A high significance level is represented by a low probability or alpha. The fact that the alpha decreases shows that there was an increasing difference between the experimental and control groups. High F-values suggest no difference in response variance between groups.

Table 4.2.1: Statistical Analysis Summary for Motivation

	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5
t-Test Figure 4.2.1a Group Analysis	0.22	0.02	0.02	0.01	0.00
F-Test Figure 4.2.1b Group Analysis	0.18	.01	.02	.78	.64



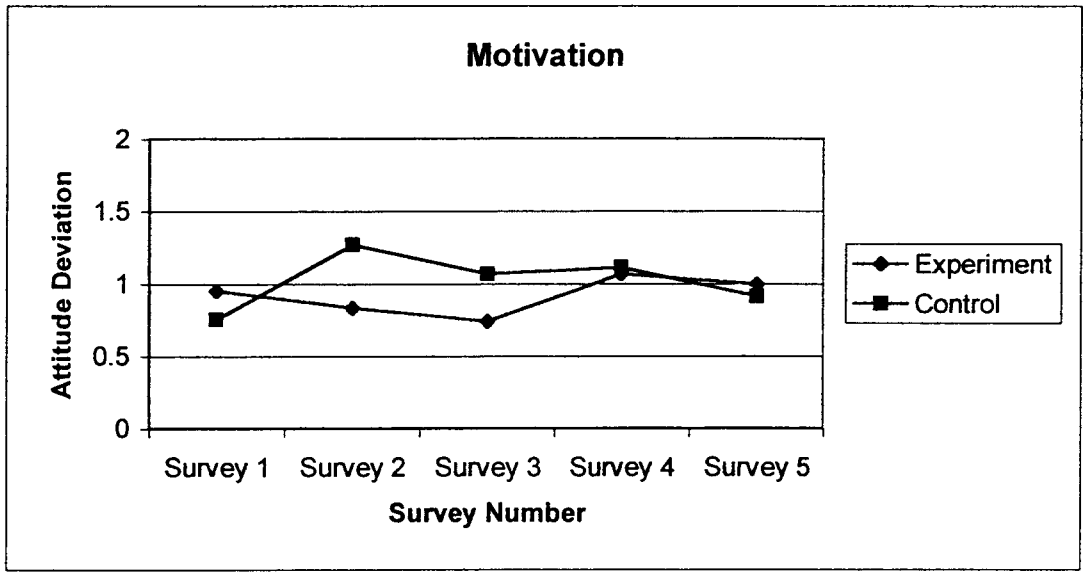


Figure 4.2.1 Survey Group: Motivation; a) mean response, b) response deviation

Table 4.2.2: Qualitative Response from Survey Question Three

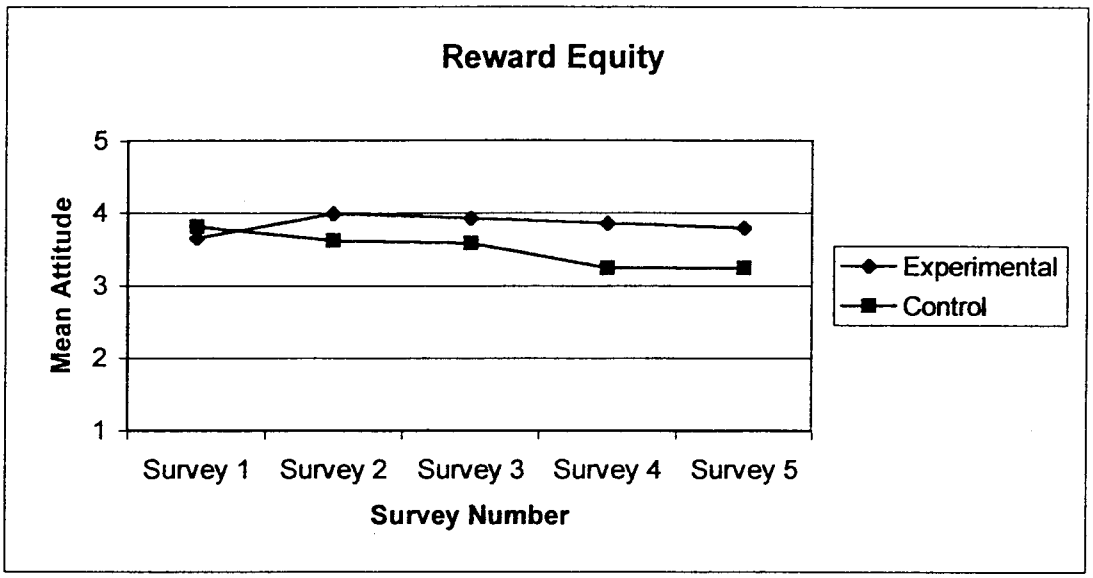
	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5
Experimental Group		Going well, feel a part	Good Discussions	Going well, feel a part of class	Class is great
		Lots of learning	Small group interactions	Very interactive	Easy to participate
		High Participation	High Participation	I contribute and learn from class	Liked attending class
		Good reading	More relaxed now	Lots of opportunity to participate	Feel comfortable in class
Positives	**N/A To this Study**	Good small and large group discussion	Higher learning	Improving over time	Class was informative
		Fast pace	Developed routine	Class increased my confidence.	
		More freedom	Presentations		
		Common interests	Improving with time		
		Good Structure	Sticking with plan		
		Feel responsible to contribute			Not enough feedback
			Holding back after being slapped down	Not enough time	Crunched at the end
		Want more from others		Didn't like evaluation	Lots of reading
Negatives		Lacks direction	Too much reading	Dislike point counting	Not enough time for abstracts
		Somewhat repetitive	Don't know grade	Need more feedback	
		Too much reading	People dominate discussion	Don't get to participate as much as I want to	Occasionally got lost in information
			Want more feedback	I feel like I'm behind	
Control Group		Interesting issues, conductive to class discussion	More group discussion	Group projects were a good way to incorporate ideas	Good Dynamics, Good interaction
		Very informative	Interactive, interesting	Good presentations	Good Speakers
		Going well	Others want my input	Class going well	It is going well
Positives	**N/A To this Study**	Information is applicable	Better than expected	Great	It's fine
		Some guest speakers are better than others	Writing and e-mail is good	Interesting	Great, last day
		Interdisciplinary is good			
		Don't feel free to be part of class	Some muddy points	Repetitious	Waste of time
Negatives		Lack direction on projects	Could be going better	Assignments not useful	Somewhat repetitive
		Too much theory	Lacks direction	Walking through a snow storm	Room for Improvement
			Slow and confusing	Homework not beneficial	Lack of participation
				Class too long	Needed more structure

Table 4.2.2 contains a summary of the comments made on the open-ended question in the motivation category. It contains both positive and negative comments about the class and how it pertains to their perception of the class. Response rate was very low on the qualitative questions. Answers were usually a single word: yes or no. The answers were consistent across time. The students on the whole liked the structure and interaction designed into the class. They wanted more feedback on a regular basis. A majority of the responses were positive.

4.2.2 Reward Equity

Reward equity is the perceived match between how well an individual feels they performed, and how they were rewarded for that performance. A high level of reward equity should yield a higher level of satisfaction in the process. The intervention or use of the balanced scorecard should have resulted in a higher level of reward equity. The balanced scorecard should better assess actual performance.

Figure 4.2.2 summarizes the survey results in the category of reward equity. There was no quantitative evidence from the survey to suggest that there was a significant difference in satisfaction between the experimental and control group. The results are summarized in Table 4.2.3. There was a higher level of significance found on survey 3 than on the other surveys. This point was prior to the midterm review in the experimental course.



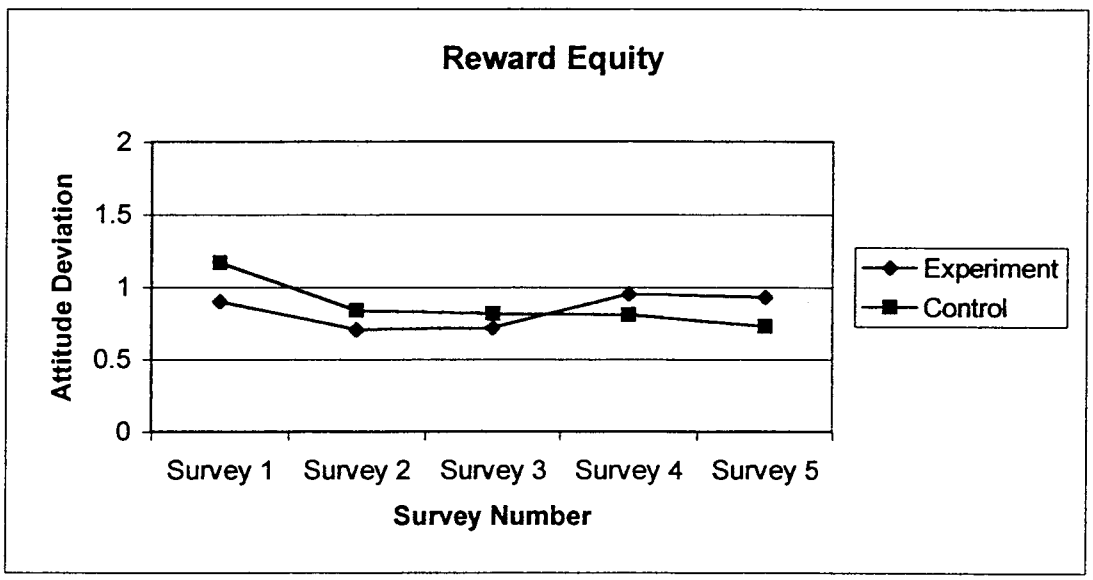


Figure 4.2.2 Survey Group: Reward Equity; a) mean response, b) response deviation

Table 4.2.3: Statistical Analysis Summary for Reward Equity

	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5
t-Test Figure 4.2.2a Group Analysis	0.75	0.56	0.13	0.35	0.70
F-Test Figure 4.2.2b Group Analysis	0.04	0.19	0.31	0.26	0.08

F-Test results show that there wasn't a significant difference in the way subjects responded to questions in during surveys 2, 3, and 4, but did respond differently on the first and last survey.

Table 4.2.4 is a summary of the comments made on Question Seven of the survey. This question targeted sources of frustration for the students. The primary sources of frustration involved the measurement of participation, lack of timely feedback, and mismatch of goals. This helps to explain why we do not see more significance between the experimental and control group in the reward equity category.

4.2.3 Satisfaction

Satisfaction is an end-of-process indicator of performance. Because performance motivation is a dynamic process, satisfaction from one experience is carried over to future situations and becomes a component of future motivation. Again, over the duration of the experiment, our hypothesis predicts that satisfaction level will increase.

Table 4.2.4: Qualitative Response from Survey Question Seven

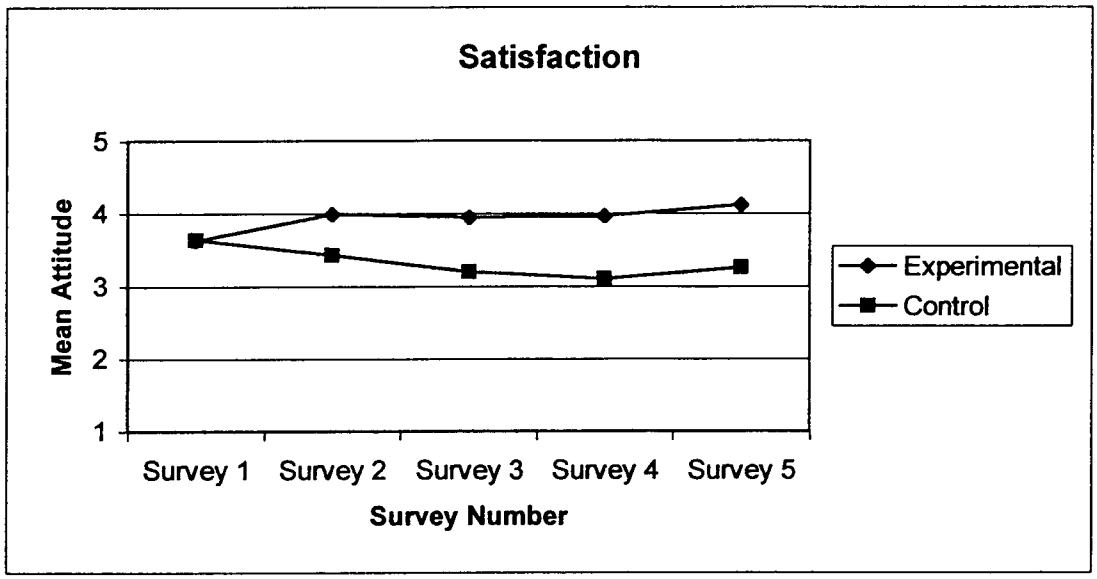
	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5
Experimental Group **Positives**		Too much reading Too fast Clear objectives Reasonable workload	So far so good No frustrations	Not really Working towards goals High participation Helped me improve Feel comfortable with structure Uninhibited discussion Increased research skills	High participation Not frustrated Helped to build confidence Learning via articles
Negatives	**N/A To this Study**	Unclear goal Not enough feedback English skills Too much group work	Scorecard decision has me worried Not enough time for presentations Participation requirements are not realistic Unclear goal Hard to communicate Groups are slow No enough feedback Seems repetitive Student presentations are not good learning tools	Problems w/ English Lots of people don't participate Worried about performance measurement Too much reading Not enough feedback Midterm review should be earlier Not enough time to correct based on midterm review	English is frustrating Mismatched goals Not enough feedback Frustrated with participation counting Sometimes hard to participate
Control Group **Positives**		No	Looking at subject matter differently, not frustrated	No time to meet in groups, scheduling time	Group projects were frustrating Changing assignments
Negatives	**N/A To this Study**	Don't know expectations Didn't like research Unclear goals Limited involvement due to time	No clear goal	Unclear goals	Unclear objectives

Figure 4.2.3 summarizes the survey results for questions targeting student satisfaction. On all but the first survey, there was a statistically significant difference between the experimental and control groups. The results from the statistical analysis are summarized in Table 4.2.5. In theory, the high level of satisfaction should carry over to future motivation. Not surprisingly, in this study both the satisfaction construct and motivation construct showed higher levels in the experimental group than in the control group.

Table 4.2.5: Statistical Analysis Summary for Satisfaction

	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5
t-Test Figure 4.2.12a Group Analysis	0.97	0.01	0.00	0.00	0.00
F-Test Figure 4.2.12b Group Analysis	0.36	0.18	0.03	0.19	0.23

A 99% significance level was realized for the difference between the experimental and control group on the satisfaction component. This supports the conclusion that satisfaction in the experimental group was positively higher than in the control group. F-Test results show that there was some difference in the way subjects responded on survey 3, but overall, there wasn't significant evidence to suggest that the groups responded differently to questions.



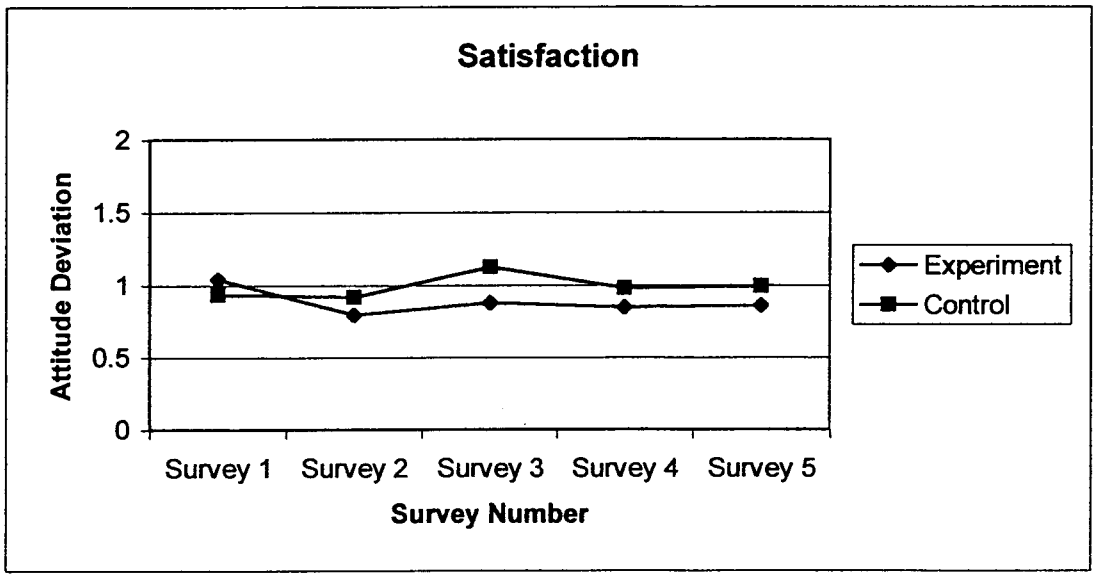


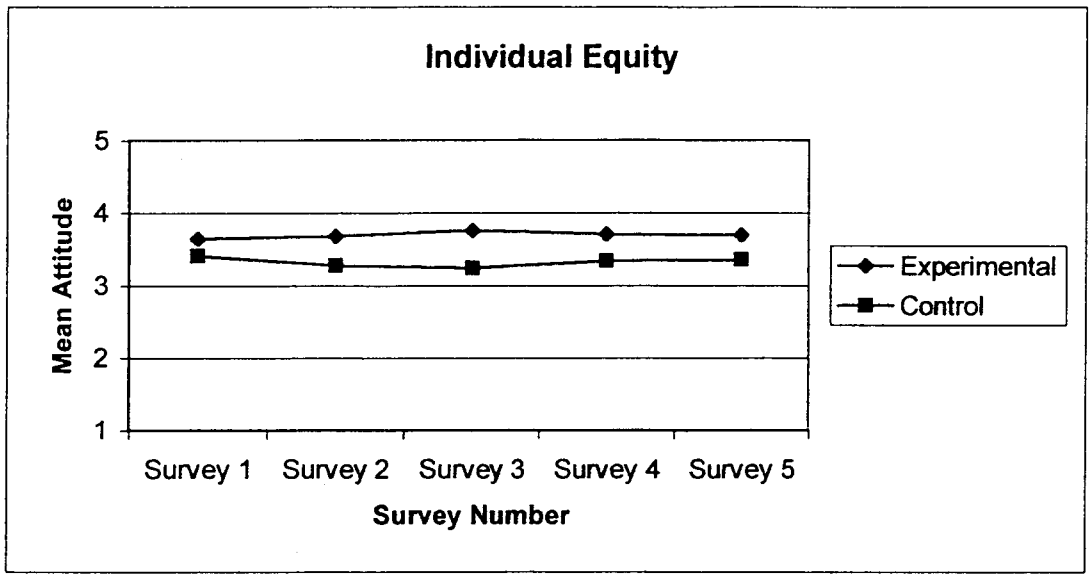
Figure 4.2.3 Survey Group: Satisfaction; a) mean response, b) response deviation

4.2.4 Individual Equity

The final group of questions addressed the issue of individual equity. Equity is how well students feel they compare with other students in the same course. For members of any group, there is a sense of cohesion or identification with others. In a classroom, all the students become members of a common group. The balanced nature of the assessment methods in the experimental group should increase individual equity because everyone is graded equally across a more balanced set of measures. This provides more opportunity for students to perform within their own skill set.

Figure 4.2.4 summarizes the survey results on the questions targeting individual equity. At the outset of this research, it was expected that there would be an approximately equal number of students in each course. Because one course had a smaller number of students enrolled in the course, it is very difficult to interpret the results on this question. In smaller groups, there is tendency to form more cohesive bonds. The results from the statistical analysis are shown in Table 4.2.6.

Unlike the first three groups, there was a split in reaction with this group. On Survey Question 14, we see that there was a significant difference between the experimental and control group. This question involves equity in class contribution. In the experimental group, there was significant evidence to suggest that there was a higher level of satisfaction of individual performance when compared with others. The assessment method in the experimental group



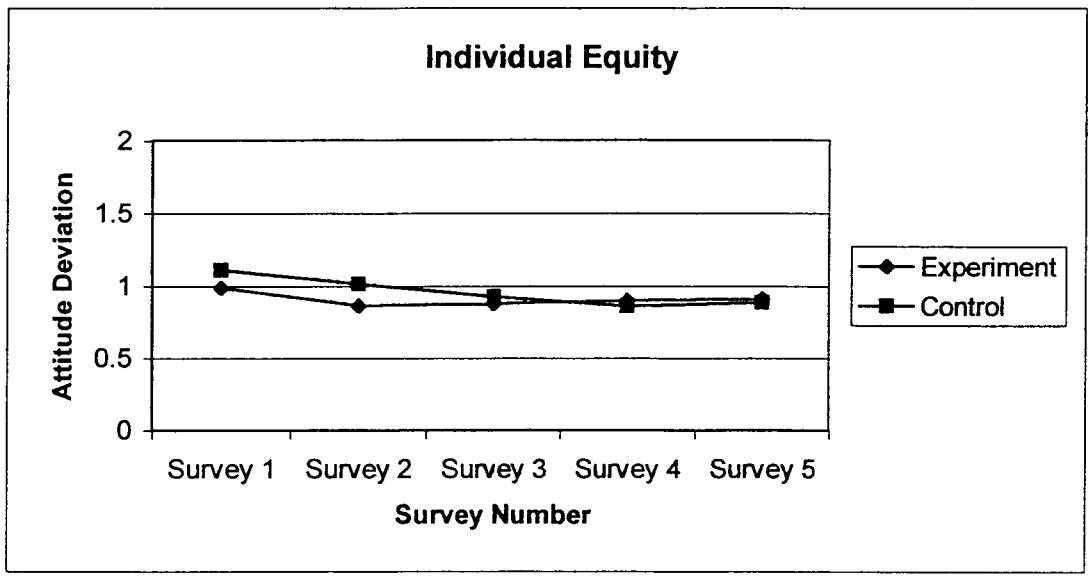


Figure 4.2.4 Survey Group: Individual Equity; a) mean response, b) response deviation

measured, and therefore valued, class contribution. Thus, it is reasonable to expect students to contribute more to discussion.

In the overall group analysis for individual equity, there was marginal evidence to suggest that individual equity was higher in the experimental group. The highest level of significance was on survey 3 (alpha = 0.02 or 98% significance). The significance range for survey 2 through 5 is 90% to 98%. This is a less significant level compared with motivation and satisfaction, but is a significant enough result to be considered important in this study.

Table 4.2.6: Statistical Analysis Summary for Individual Equity

	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5
t-Test Figure 4.2.4a Group Analysis	0.35	0.08	0.02	0.07	0.10
F-Test Figure 4.2.4b Group Analysis	0.30	0.14	0.62	0.72	0.84

4.2.5 Inter-Item Reliability

When performing research using a survey methodology, it is important to determine the reliability of the survey instrument to measure the desired phenomenon. Using multiple survey questions to address one construct is one method of increasing survey reliability. An inter-item reliability test can be performed to find the correlation level between items. This test is usually performed using Cronbach's Alpha test. Table 4.3.7 summarizes the result of the inter-item correlation between the groups of survey questions. In social research,

correlation of .35 to .5 are considered to be high (Miller, 1991). As can be seen, in the five times the survey was issued, the inter-item correlation ranged from .42 to .91 which is high for social research and thus provides and shows a high level of repeatability within the survey, and increased validity for the measured items.

Table 4.2.7: Group Inter-Item Correlation Level

	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5
Group 1 Motivation	.46	.74	.77	.76	.83
Group 2 Reward Equity	.62	.54	.76	.77	.80
Group 3 Satisfaction	.81	.77	.87	.85	.91
Group 4 Individual Equity	.42	.58	.78	.72	.77

4.3 Direct Observation

Direct observation was performed to measure student participation in the class. Participation and contribution to discussion were measured as part of the balanced scorecard intervention and therefore deemed valuable by the students. The participation metric was suggested by the students in the course. Inclusion of the participation metric in the balanced scorecard provided a sense of ownership of the measurement system by the students in the experimental course.

Participation data can be seen in Figure 4.3.1. The data is standardized and presented in number of comments per hour of class per week. Because the two

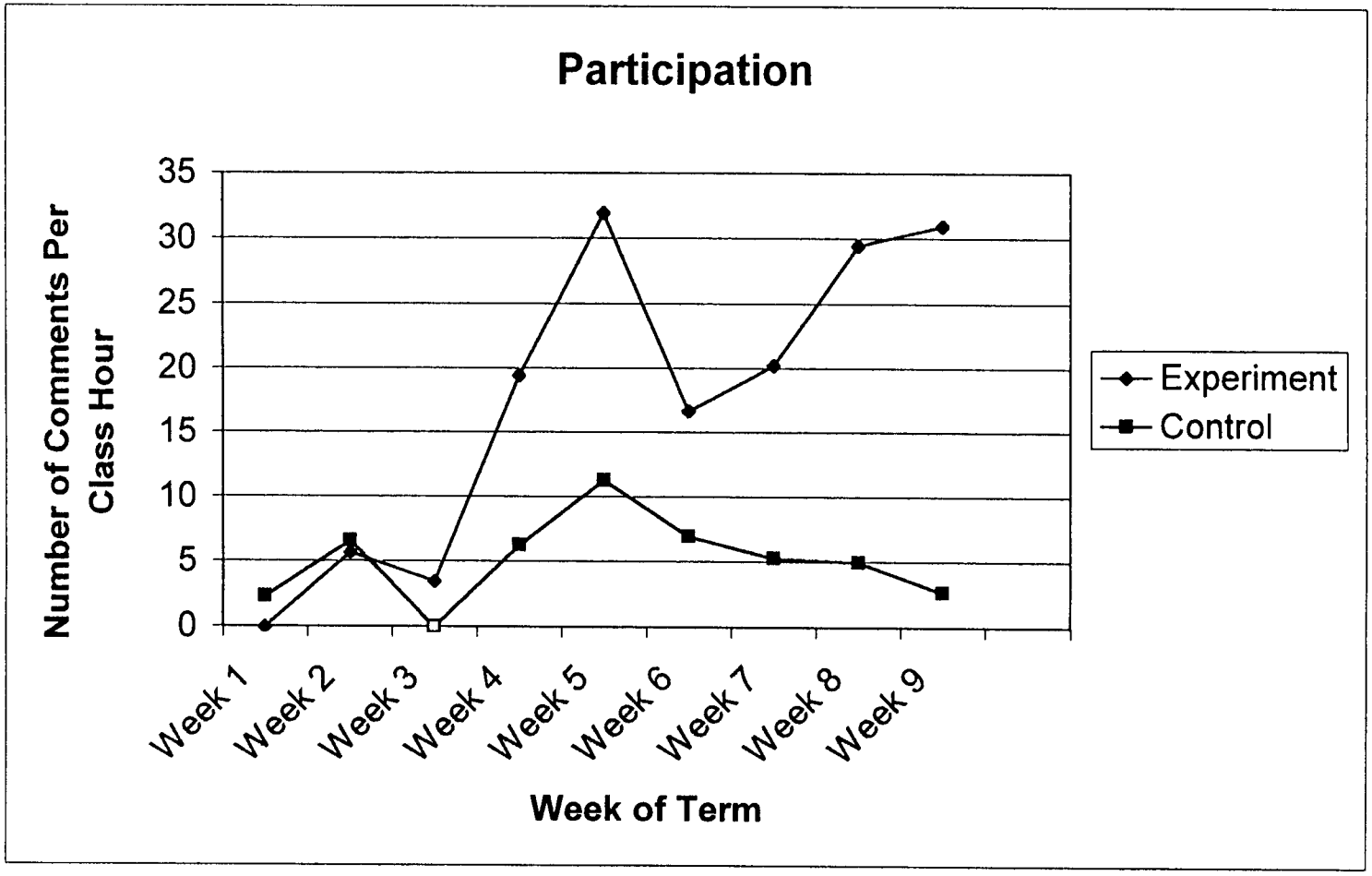


Figure 4.3.1 Participation Comparison

classes had a different amount of scheduled time, the number of comments could not be directly compared without the use of a conversion.

The experimental group had an increasing level of participation throughout the term. There was a spike in the participation for the experimental group during week 5 due to a highly interactive guest lecturer. In the control group, there was no participation data recorded during week 3 because there was no class held on that week. A t-test for the average participation over the course of the term showed that the participation in the experimental group was higher than in the control group with a significance level of 99%.

The participation data has two important impacts. First, participation increased during the duration of the term. This suggests that as the term progressed something encouraged students to participate or contribute more in class discussion in the experimental course. This directly impacts the results of this study. Second, because at the foundation of this research is the desire to engage students in the educational process, it has been shown that this intervention was able better able to connect students with the course.

4.4 Focus Groups

The intent behind the focus groups was to gain a better understanding of the factors driving student attitudes in the experimental class. The qualitative data collected would be used to understand the factors that drove the student's actions during the term. The information was collected at the conclusion of survey and

direct observation data collection, and prior to the release of grade information to prevent biasing the results.

Table 4.4.1 summarizes the triangulated results from the focus group sessions. The results are grouped by question and basic themes including administration, participation, and midterm review. There were both positive and negative feedback about the balanced scorecard approach collected in the focus groups. Strengths weaknesses could also be identified through the focus group process.

Positive aspects included the midterm evaluation, and the balanced nature of the metrics used in the balanced scorecard assessment. The students felt that the continual collection of performance data emphasized learning over passing tests and reduced performance stress. One student commented "It's more like what we would see in the real world," referring to performance assessment in the workplace.

Negative aspects included the subjective nature of some of the balanced scorecard measurements and some of the logistics involved with implementing it. Students were frustrated with the implementation of the balanced scorecard because they were used to having the grading system laid out clearly at the beginning of the term. They were also concerned with the subjective nature of some of the measures.

The strengths of the system were the feeling of fairness and the high feedback provided to the students. While the students felt that the feedback may not have occurred when and as often as they would have liked, they did feel that the feedback that they received was very valuable. They felt that the balance of

Table 4.4.1: Triangulated Focus Group Results

	Administration	Participation	Midterm Review
What have you disliked about the Balanced Scorecard assessment method?	Vagueness No clear Definition from outset More than one person grading	Counting of participation Group participation not included Over participation Personality not taken into account	Occurred too late No opportunity to improve
What have you liked about the Balanced Scorecard assessment method?	Student Ownership Student Self Evaluation Student input on what gets graded		Balance of Measures Learning vs. Testing emphasis Feedback Communication between professor and student
How do you feel the Balanced Scorecard compares with traditional assessment methods?	*This question compared traditional and the BSA assessment technique:* BSA is more subjective Requires more effort on part of the professors BSA emphasizes learning		
Would you want to take another course that utilized a Balanced Scorecard assessment method?	*The answer was primarily yes with reservation including:* Course content Professor(s) Selected metrics for the Balanced Scorecard More feedback sooner		
Did you feel that the assessment methods were fair?	Balance makes it more fair	Didn't like having participation counted (subjective) Time constraints in class Comment Quality	Self-assessments valued
How did the assessment measures in this class effect your motivation?	Timing of the instruction and assessment outline	High level of participation as a result from the BSA	Timing of the review, number of the reviews Feedback occurred too late Feedback increased participation
At any point during the term, were you frustrated with the way you were assessed in this class? Why?	Implementation was unclear and not fully understood.	Surprised by participation results High value of participation	Mismatch between self and professor evaluation Feedback timing and number of reviews
Is there anything else that you would like to add about your experience with the assessment method this term?	Need for metric improvement Learning vs. Passing tests.	No lumpy accumulation of points Student mood, personality, and culture effects participation	More reviews Midterm review excellent feedback source

measures made it fairer than traditional grading. All three groups felt that they would be willing to take another course that utilized the balanced scorecard. Student ownership was also viewed a strength of the balanced scorecard assessment. One student said "I felt more connected to the course".

The weaknesses that can be drawn deal with timing and logistics. Since this was the first time that the balanced scorecard was being used in a classroom environment, there were some unplanned problems that had to be dealt with. The use of the midterm review was difficult to plan and orchestrate. It took a great deal of planning to prepare the student evaluations on the part of both the students and professors. The midterm occurred late in the term as a result. Also, because student input was utilized to design the assessment metrics, the balanced scorecard wasn't finalized until the third week of the term. This caused frustration on the part of the students.

4.5 Data Triangulation

This study used data from multiple sources in order to understand student attitude and reactions to the balanced scorecard intervention during the course of the term. This section cross-examines the data from the three data sources.

From the survey, we found that student motivation and satisfaction increased over the duration of the term in the experimental course. Participation data from direct observation was also found to be significantly larger in the experimental course. This is a logical result of higher motivation to perform. Because

performance motivation is a dynamic process, increased satisfaction led to increased motivation, followed by an increase in participation. A majority of the subjects in the focus group suggests that students liked the balanced scorecard assessment method better than traditional methods. Subjects stated that they would like be willing to take another course that utilized a balanced scorecard assessment approach. This shows a high level of satisfaction with the process, and a desire or motivation to be involved with a possible future course.

The survey results show that the individual equity was moderately better in the experimental course. The high participation suggests that the subjects participated more in interactive discussion. This creates a personal connection between students, and gives them the chance to express their views. This type of activity can increase individual equity. The focus group provided more support for these conclusions. The balance of measurements used in the experimental group provided more opportunities for individuals to contribute and be a part of the class. Including students in the process of creating and selecting measures also gave the students a sense of ownership. The group ownership also increases individual equity. Working against individual equity is the subjectivity of some of the measures. It contributed to some students feeling that they were unfairly evaluated.

There was no significant difference found between the experimental and control groups with the survey in regard to reward equity. This can be explained by the primary problems identified in the focus groups. The timing and administrations issues are probably the most significant. Students felt unclear as to

expectations and to how they were being evaluated because of the late implementation (Week 3) of the balanced scorecard. The midterm review also occurred late in the term. During the review process, some subjects experienced a mismatch between the student self-evaluation and the professor evaluation. This can lead to a reduction in reward equity. One student stated,"The midterm review occurred at the end of the term, and there wasn't a chance to fix it (the evaluation)". Disconnects between expectation and the evaluation would result in a reduction in reward equity. The important thing to remember is that the problems didn't reduce reward equity below the control group, as would be indicated by a significant negative result.

Chapter 5: Conclusion

5.1 Chapter Overview

This chapter brings together theory and research findings to be able to understand how the balanced scorecard intervention effected student motivation. Also discussed in this chapter are the lessons learned while using the balanced scorecard as an educational assessment tool and the direction that future research should take.

5.2 What Was Found

This research began with the question "Can the use of a broad-based holistic assessment method positively affect student motivation?" The results from this research provide sufficient evidence to suggest that the use of a broad-based holistic assessment method can increase student motivation. Three out of the four motivation components studied showed marginally to significantly higher levels under the balanced scorecard assessment technique. Participation also increased showing engagement of the students in the class.

During the focus group sessions, learning versus completing assignments was highlighted as an important difference between traditional assessment

techniques and the balanced scorecard technique. This further demonstrates the ability to engage students with the content and structure of the course.

Engaging students is a motivation problem. Low motivation yields student apathy. To engage students requires high levels of motivation. Students must feel connected with the subject of the course, but also must have a sense that they can succeed. The performance assessment system is the structure that can support or restrict student performance. Therefore, in order to create and maintain an environment of high motivation and student engagement, an appropriately designed performance assessment is vital. To accomplish this the human factors of assessment and feedback should be leveraged to engage the student constructively.

Motivation is a dynamic process. Evidence suggests that the level of motivation was not stagnant. Student motivation was in a state of flux, reacting to changes in the classroom environment, particularly in regard to timing and feedback, key features of effective performance assessment systems. The logistics involved with using the balanced scorecard assessment approach caused some problems. Specifically, the focus groups identified that timing and feedback (midterm reviews) were a major factor contributing to frustration during the term.

This section should include student reaction to the balanced scorecard assessment. Some important benefits identified by students were student ownership, balanced fairness of the assessment, and student self-evaluation as a component of the overall assessment. Even more important was the strongly positive response by students when asked if they would want to take another course

graded with a balanced scorecard. Even though many weaknesses were identified, the positive response suggests a high level of satisfaction as a result of being assessed with the balanced scorecard.

Designing performance assessment systems is incredibly challenging. Assessments fulfill a variety of different intents and purposes. While performance evaluation and ranking components are important, they are not more important than the motivational component. This research, if nothing else, shows that the motivational component is extremely valuable and should be considered in the design of courses and performance assessment methods.

5.3 What Should Have Been Done Differently

The focus groups identified several weaknesses or deficiencies with the balanced scorecard intervention, primarily the timing and amount of feedback. The balanced scorecard assessment method by design is intended to be a highly interactive assessment technique. The intent of the midterm review was to provide an interactive feedback mechanism between students and instructors. While on the whole, the midterm review was a valuable process, the feedback that was gained by them was deemed too late.

Feedback needs to be delivered in a timely manner, and be connected tightly to performance. To accomplish this, there should be more than one midterm review. Only having the one major point of contact concentrates feedback. There needs to be a regular and continuous feedback process.

The requirements need to be laid out earlier in the term. While it would be impossible to get student input before the first class meeting, and effort should be put forth to collect student input and finalize measures in the first week. In addition, sometime during the middle of the term, the measures should be revisited to see if revisions are needed. If a measure is deemed inappropriate or ineffective, it can be revised or eliminated to better match student and instructor needs.

5.4 Areas of Future Research

Performance measurement is not an exact science. This research has presented a different approach to performance measurement. Because this method departed from traditional techniques, new metrics were required to evaluate performance. The metrics were developed collaboratively between students and instructors. It is unlikely that the metrics were optimally selected to evaluate performance. One potential area for future work is the development of measurement guidelines. The guidelines could help instructors align course objectives with appropriate evaluation metrics.

This research looked at a specific student population. The model of motivation dynamics can not be generalized to most populations. The balanced scorecard assessment is a specific type of organizational support that targeted a specific population. To increase the generalizability of the findings, there is a need to apply the balanced scorecard to a variety of settings and look at the motivation effects. The problem with most research in the area of measurement is the

concentration of performance improvement versus motivation improvement. Performance is a direct function of motivation (Lawler, 1994). Therefore, increasing motivation should improve performance.

This research could also be extended into the business world. Motivation is often overlooked in most organizations. Yet, motivation is a major driver behind most work (Lawler, 1994). Many firms have adopted a balanced scorecard approach to measurement. An emphasis is placed on performance improvement versus motivation improvement. The balanced scorecard has seen a lot of success though. Could the success seen by these companies be correlated back to improvements in motivation?

In the past few years the introduction of Certificates of Initial and Advanced Mastery (CIMs and CAMs), and the use of portfolios for assessment have changed the way in which performance is evaluated. The intent behind these techniques is to broaden the range of measurements being used to evaluate performance. This is similar to the concepts explored in the balanced scorecard assessment method. Unfortunately, they do not take into account student motivation as part of the assessment. Because performance is a function of motivation, there is a need to focus on engaging and motivating students in the learning process rather than trying to directly improve student performance.

5.5 Concluding Remarks

The world is quickly moving into an information age that is going to require individuals that are highly motivated and creative. Unfortunately, motivation is not just about people. It is about the dynamic interaction between people and the organizational structure that supports them. This is true whether the organization is a business, and educational institution, or a volunteer group. People need to have direction, support, and need appropriate recognition for their performance.

This research focused on performance assessment in education and its effect on the dynamic motivation cycle. Performance assessment is a component of the organizational support mechanisms in place to help students learn. Most performance assessment tools currently in place do not take the dynamic motivation cycle into consideration in the assessment cycle. They only look at performance snap shots in time to evaluate performance, and do not support a continuous process of holistic and motivational assessment.

The alternative performance assessment method proposed in this research is called the balanced scorecard. The balanced scorecard, unlike most assessment techniques, takes a more holistic approach to measurement. The contention of the researcher is that by incorporating a balanced scorecard approach to educational assessment, an increase or positive escalation in motivation could be realized. Motivation is the driving force behind performance. Without considering motivation in the performance picture, you can never expect to engage a person's full capability.

So, what is an A you ask? It is more than ranking. It is more than an incentive. It is more than a milestone. It represents a system of all these things and more. An A is a symbol with causes and consequences. Until people understand the full assessment picture, they can never hope to truly appreciate what assessment does or is potentially capable of doing.

Bibliography

Aken, E. M. V. (1991). A multiple case study on the information system to support self-managing teams. Unpublished master's thesis, Virginia Polytechnic Institute and State University, Blacksburg.

Amabile, T. M. (1997). Motivation creativity in organizations: On doing what you love and loving what you do. California Management Review, 40(1), 40-58.

American Society for Engineering Education. (1998). How do you measure success?: Designing effective processes for assessing engineering education. Washington, DC: ASEE Professional Books.

Amidon, D. M., & Skyrme D. J., (1998). New Measures of Success. Journal of Business Strategy, 19(1), 20-24.

Angelo, T. A, & Cross, P. K. (1993). Classroom Assessment Techniques: A Handbook for College Teachers. San Francisco, CA: Jossey-Bass Publishers.

Argyris, C. (1994). Good communication that blocks learning. Harvard Business Review, 72(4), 77-85.

Azwell, T. & Schmar, E. (1995). Report card on report cards: Alternatives to consider. Portsmouth, NH: Heinemann.

Babcock, D. L., (1996). Managing Engineering and Technology. Upper Saddle River, New Jersey: Prentice Hall.

Blumberg, M. & Pringle, C.D. (1982). The missing opportunity in organizational research: some implications for a theory of work performance. Academy of Management Review, 7(4), 560-569.

Bruns, W. J., Jr. (1992). Performance, measurement, evaluation, and incentives. Boston, MA: Harvard Business School Press.

Byham, W. C., Cox, J. C., & Shomo, K. H. (1992). Zapp! in education: How empowerment can improve the quality of instruction, and student and teacher satisfaction. New York, NY: Development Dimensions International, Inc.

Chow, C. W., Haddad, K. M., & Williamson, J. E. (1997). Applying the balanced scorecard to small companies. Management Accounting, 78(8), 21-27.

Creswell, J. W. (1994). Research design: Qualitative & quantitative approaches. Thousand Oaks, CA: Sage Publications, Inc.

Deming, W. E. (1986). Out of the crisis. Cambridge, MA: Massachusetts Institute of Technology, Center for Advanced Engineering Study.

Deming, W. E. (1994). The new economics. Cambridge, MA: Massachusetts Institute of Technology, Center for Advanced Engineering Study.

DeVries, D. L., Morrison, A. M., Shullman, S. L. & Gerlach, M. L. (1981). Performance Appraisals on the line. New York, NY: John Wiley & Sons, Inc.

Fink, A., Bourque, L. B., Fielder, E. P., Frey, J. H., Oishi, S. M., & Litwin, M. S. (1995). The Survey Kit. (Vols. 1-9). Thousand Oaks, CA: Sage Publications, Inc.

Fowler, F. J. Jr. (1988). Survey Research Methods. Newbury Park, CA: Sage Publishers, Inc.

Frey, B. S. (1997). On the relationship between intrinsic and extrinsic work motivation. International Journal of Industrial Organization, 15(4), 427-439.

Frohman, A. L. (1998). Igniting organizational change from below: The power of personal initiative. Organizational Dynamics, 26(1), 39-53.

Greve, H. R. (1998). Performance, aspirations, and risky organizational change. Administrative Science Quarterly, 43(1), 58-86.

Hansen, D. G. (1997). Worker performance and group incentives: a case study. Industrial and Labor Relations Review, 51(1), 37-49.

Ittner, C.0 D., & Larcker, D. F. (1997). The performance effects of process management techniques. Management Science, 43(4), 522-534.

Jenkins, L. (1997). Improving Student Learning. Milwaukee, WI: ASQC Quality Press.

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard – measures that drive performance. Harvard Business Review, 70(1), 71-79.

Kaplan, R. S., & Norton, D. P. (1993). Putting the balanced scorecard to work. Harvard Business Review, 71(5), 134-147.

Kaplan, R. S., & Norton, D. P. (1996a). The Balanced Scorecard. Boston. Massachusetts: Harvard Business School Press.

Kaplan, R. S., & Norton, D. P. (1996b). Linking the Balanced Scorecard to Strategy. California Management Review, 39(1), 53-79.

Kaplan, R. S., & Norton, D. P. (1996c). Using the Balanced Scorecard as a Strategic Management System. Harvard Business Review, 74(1), 75-85.

Kaplan, R. S., & Norton, D. P., (1997). Why does business need a balanced scorecard? Journal of Cost Management, 11(3), 5-10.

Kerr, S. (1975). On the folly of rewarding A, while hoping for B. Academy of Management Review, 18(4), 769-783.

Kohn, A. (1993a). Why incentive plans cannot work. Harvard Business review, 71(5), 54-63.

Kohn, A. (1993b). Punished by rewards: The trouble with gold stars, incentive plans, a's, praise, and other bribes. New York, NY: Houghton Mifflin Company.

Kohn, A. (1998). What to look for in a classroom. San Francisco, CA: Jossey-Bass Publishers.

Krueger, R. A. (1994). Focus Groups: A practical guide for applied research. Thousand Oaks, CA: Sage Publications, Inc.

Lawler, E. E. (1991). High-involvement management. San Francisco, CA: Jossey-Bass Publishers.

Lawler, E. E. (1994). Motivation in Work Organizations. San Francisco: Jossey-Bass Publishers.

Longenecker, C. O., Stansfield, T. C., & Dwyer, D. J. (1997). The human side of Manufacturing Improvement. Business Horizons, 40(2), 7-17.

Lynch, R. L., & Cross, K. F. (1993). Measure up: Yardsticks for continuous improvement. Cambridge, MA: Blackwell Publishers.

McWilliams, B. (1996). The Measure of Success: In between the bottom line and employee-satisfaction surveys lies the best way to rate your company: the balanced scorecard. Across the Board, 33(2), 16-20.

Meyer, C. (1994). How the right measures help teams excel. Harvard Business Review, 72(3), 95-103.

Miller, D. C, (1991). Handbook of research design and social measurement. Newbury Park, CA: Sage Publications, Inc.

Mitchell, T. (1982). Motivation: direction for theory, research, and practice. Academy of Management Review, 71(1), 80-88.

Montgomery, D. C. (1997). Design and analysis of experiments. New York, NY: John Wiley & Sons.

Mount, M. K. (1983). Comparisons of managerial and employee satisfaction with a performance appraisal system. Personnel Psychology, 36, 99-110

Murphy, K. R., & Cleveland, J. N. (1995). Understanding performance appraisal: Social, organizational, and goal-based perspectives. Thousand Oaks, CA: Sage Publication, Inc.

Newing, R. (1995). Wake up to the balanced scorecard! Management accounting: journal of the Institute of Cost and Works Accounts, 73(3), 22-23.

Norusis, M. J. (1993). SPSS for windows, professional statistics release 6.0. Chicago, IL: SPSS, Inc.

Patton, M. Q. (1990). Qualitative Evaluation and research methods. Newbury Park, CA: Sage Publications, Inc.

Petrick, J. A., & Furr, D. S. (1995). Total quality in managing human resources. Delray Beach, FL: St. Lucie Press.

Pfeffer J. (1998). Six dangerous myths about pay. Harvard Business Review, 76(6), 109-119.

Popham, W. J. (1995). Classroom assessment: What teachers need to know. Needham Heights, MA: Allyn & Bacon, A Simon & Schuster Company.

Powell, J. (1998). How incentives undermine performance. The Journal for Quality and Participation, 76(1), 6-13.

Reineke, R. A. (1998). Challenging the mind, touching the heart: Best assessment practices. Thousand Oaks, CA: Corwin Press, INC.

Rothney, J. W. M., Danielson, P. J., & Heimann, R. A. (1959). Measurement for guidance. New York, NY: Harper & Brothers.

Rowntree, D. (1987). Assessing students: How shall we know them? New York, NY: Nichols Publishing Company.

Schermerhorn, J. R. (1992). Leading through motivation. Management for Productivity (pp.438-469). New York, NY: John Wiley & Sons, Inc.

Senge, P. M. (1990). The fifth discipline: The art of the learning organization. New York, NY: Doubleday, a division of Bantam Doubleday Dell Publishing Group, Inc.

Smither, J. W. (1998) Performance appraisal: State of the art in practice. San Francisco, CA: Jossey-Bass Publishers Inc.

Spitzer, D. R. (1995). Supermotivation. San Francisco, CA: AMACOM, a division of American Management Association.

Stewart III, G. B., Applbaum, E., Beer, M., Lebby, A.M., Amabile, T.M., McAdams, J., Kozlowski, L.D., Baker, G.P. & Wolters, D.S. (1993). Rethinking rewards. Harvard Business Review, 71(6), pp. 37-49.

Storey, A. G. (1970). The measurement of classroom learning. Chicago, IL: Science Research Associates, INC.

Svinicki, M. D. (1998). Helping students understand grades. College Teaching, 46(3), 101-105.

Walton, R. E. (1985). From control to commitment in the workplace. Harvard Business Review, 63(2), 77-84.

Weston, P. (1991). Assessment of pupil achievement: Motivation and school success. Rockland, MA: Swets & Zeitlinger INC.

APPENDICES

Appendix A: Informed Consent and Sample Survey

Department of Industrial & Manufacturing Engineering
Oregon State University

INFORMED CONSENT DOCUMENT

I understand that I will participate in research conducted under the supervision of Dr. Kimberly Douglas of the Industrial & Manufacturing Engineering Department. I understand that in this research involves studying the effects due to the introduction of a broad base evaluation system. I understand that my performance is not in question, and that my perceptions and behavior are of interest. This information will be collected by surveys, interviews, and direct observation applied at the throughout the course of this class. I understand that the total time commitment over the entire term will not amount to more than two hours.

I am aware that this study is unpaid and that I will receive no rewards or penalties as a result of participation. My responses from participation in this study will not be taken into account during course evaluation.

My identification will not be released to any other persons, organizations, or publications. All references to subjects in this study will be encoded and kept confidential, and all identity related information destroyed within three years of the experiment. I also understand that my personal responses in association with my name will be kept confidential and will not be seen by any faculty.

I understand that any questions concerning aspects or rights related to this experiment should be directed to Dr. Kimberly Douglas at 541-737-3644 or to Paul Brotherton at 541-752-6104. I understand that Oregon State University does not provide compensation or medical treatment in the event the subject is injured as a result of participation in this study.

I understand that participation is voluntary, and my refusal to participate will not result in penalties or loss of benefits to which I am otherwise entitled. My signature below indicates that I have read and that I understand the procedures described above and give my informed and voluntary consent to participate in this study. I understand that I will receive a signed copy of this consent form.

_____ _____

Subject's Signature Date Signed

_____ _____

Subject's Name Subject#

_____ _____

Subject's Phone Number

_____

_____

Subject's Address

Subject #_____

The questions on this survey are designed to be answered while you consider what has happened in class to this point. **Consider only the class in which you were given this survey.**

1. How well have you liked this class?
   Complete Satisfaction
   Moderate Satisfaction
   No Feeling
   Moderate Dislike
   Strong Dislike

2. Would you advise a friend to take this class?
   Yes
   Pro-con
   No

3. Describe how you feel this class is going? Do you feel you have been part of this class?

4. How well satisfied are you with your current class standing?
   a. Very Well Satisfied
   b. Somewhat Satisfied
   c. Neither Satisfied or Dissatisfied
   d. Somewhat Dissatisfied
   e. Very Dissatisfied

5. How well satisfied are you with your chance of improving your evaluation in this course?
   a. Very Well Satisfied
   b. Somewhat Satisfied
   c. Neither Satisfied or Dissatisfied
   d. Somewhat Dissatisfied
   e. Very Dissatisfied

6. How satisfied are you with the way things are going in this class?
   a. Very Well Satisfied
   b. Somewhat Satisfied
   c. Neither Satisfied or Dissatisfied
   d. Somewhat Dissatisfied
   e. Very Dissatisfied

7. In this course, have you felt that frustrated with your ability to progress towards your or the class's goal? Is so, describe an example.

8. How well do you like the sort of class work you were doing?
   a. Like it a lot
   b. Like it a little
   c. Don't Care
   d. Dislike it a little
   e. Dislike it a lot

9. Do you feel assignments give you a chance to do what you do best?
   a. Yes, a lot
   b. Yes, a little
   c. Not especially
   d. No, a little
   e. No, a lot

10. Do you feel a sense of accomplishment from the work you are doing?
    a. Strong Sense of Accomplishment
    b. Moderate Sense of Accomplishment
    c. Some Sense of Accomplishment
    d. A Little Sense of Accomplishment
    e. No Sense of Accomplishment

11. I feel my contribution in this course is (of) _____?
    a. High Importance
    b. Moderate Importance
    c. Some Importance
    d. Little Importance
    e. No Importance

12. How well do you think you compare with others in the class at getting things done?
    a. Very Good, Best in Class
    b. Good, High in Class
    c. Okay, Average in Class
    d. Poor, Low in Class
    e. Very Poor, Worst in Class

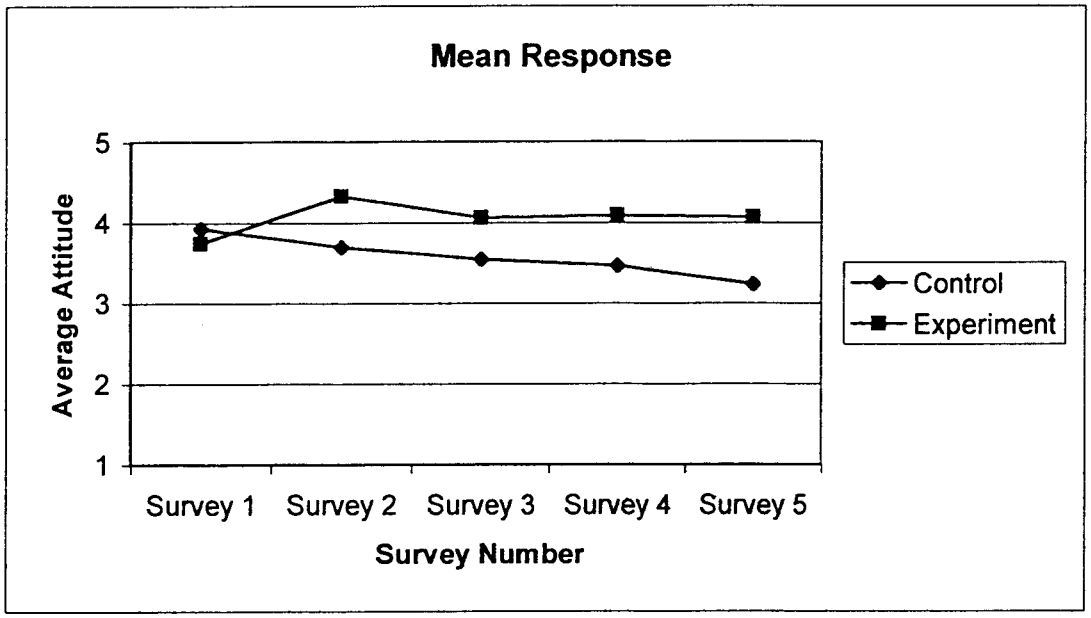
13. How well do you think you compare with others in the class in quality of work?
    f. Very Good, Best in Class
    g. Good, High in Class
    h. Okay, Average in Class
    i. Poor, Low in Class
    j. Very Poor, Worst in Class

14. How well do you think you compare with others in the class in contribution to discussion?
    k. Very Good, Best in Class
    l. Good, High in Class
    m. Okay, Average in Class
    n. Poor, Low in Class
    o. Very Poor, Worst in Class

15. What level of identification do you feel with others in the class?
    a. Strong Identification
    b. Moderate Identification
    c. Weak or Lack of Identification

Appendix B: Motivation Results

a)



b)

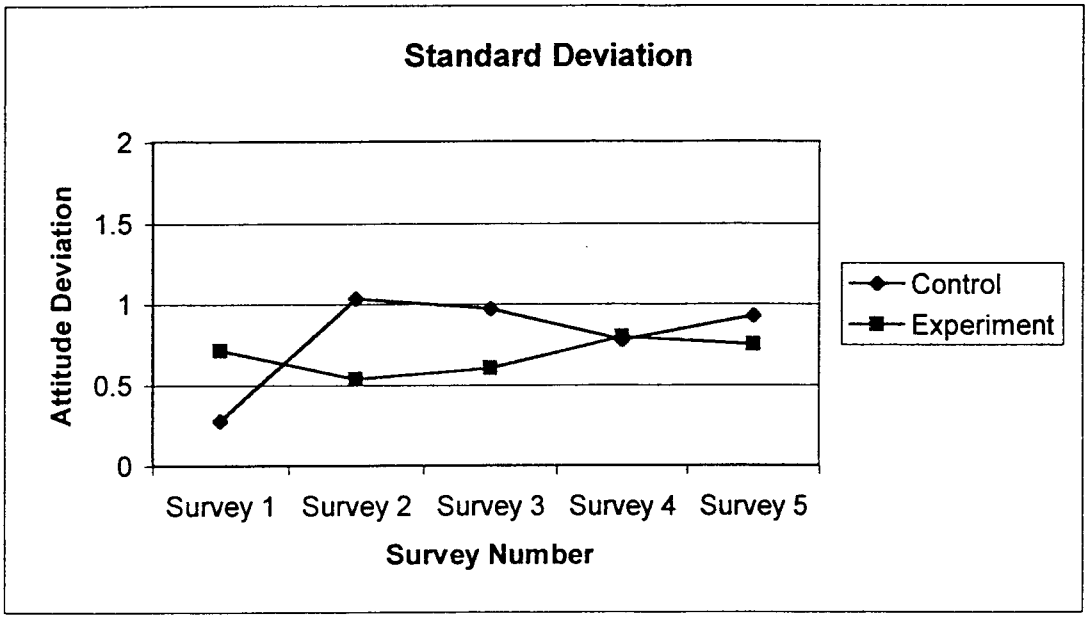
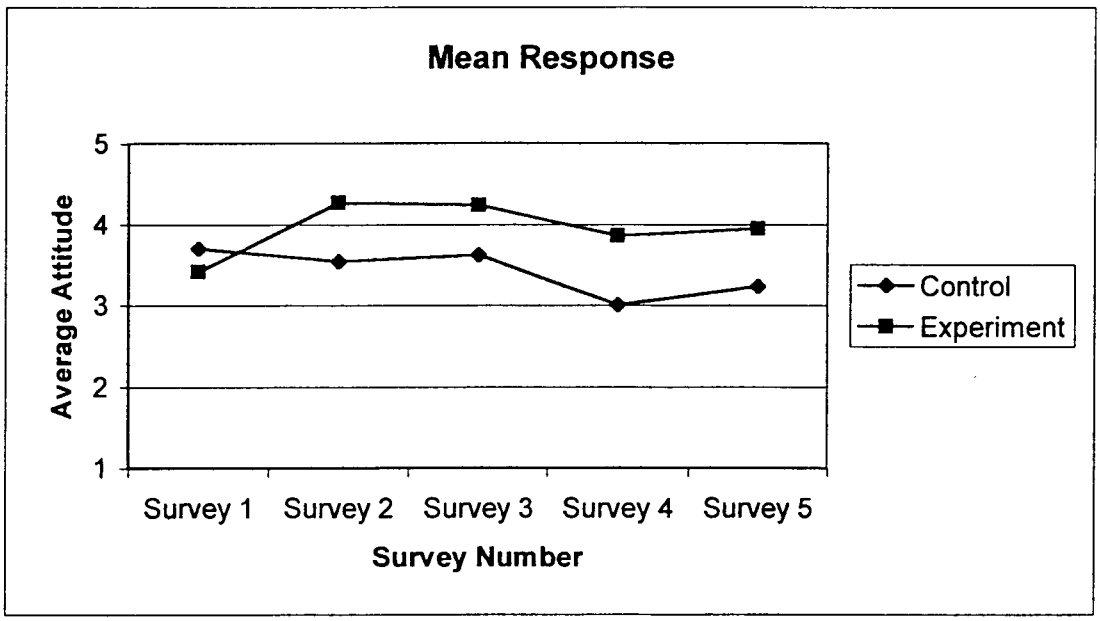


Figure B.1 Survey Question: How well have you liked this class?; a) mean response, b) response deviation

a)



b)

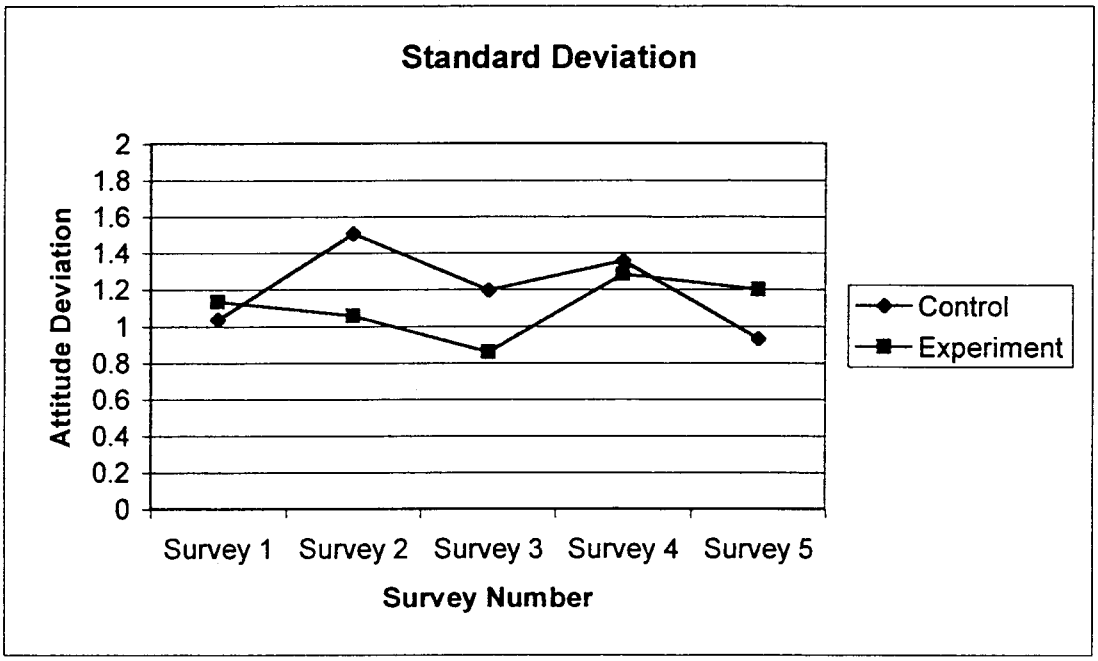
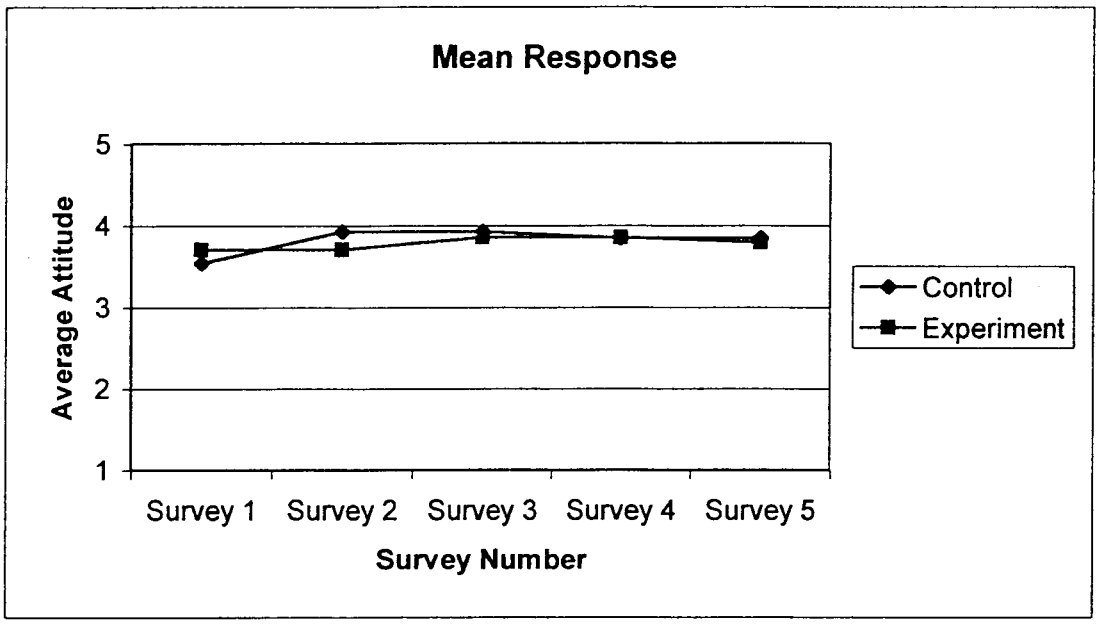


Figure B.2 Survey Question: Would you advise a friend to take this course?; a) mean response, b) response deviation

Appendix C: Reward Equity Results

a)



b)

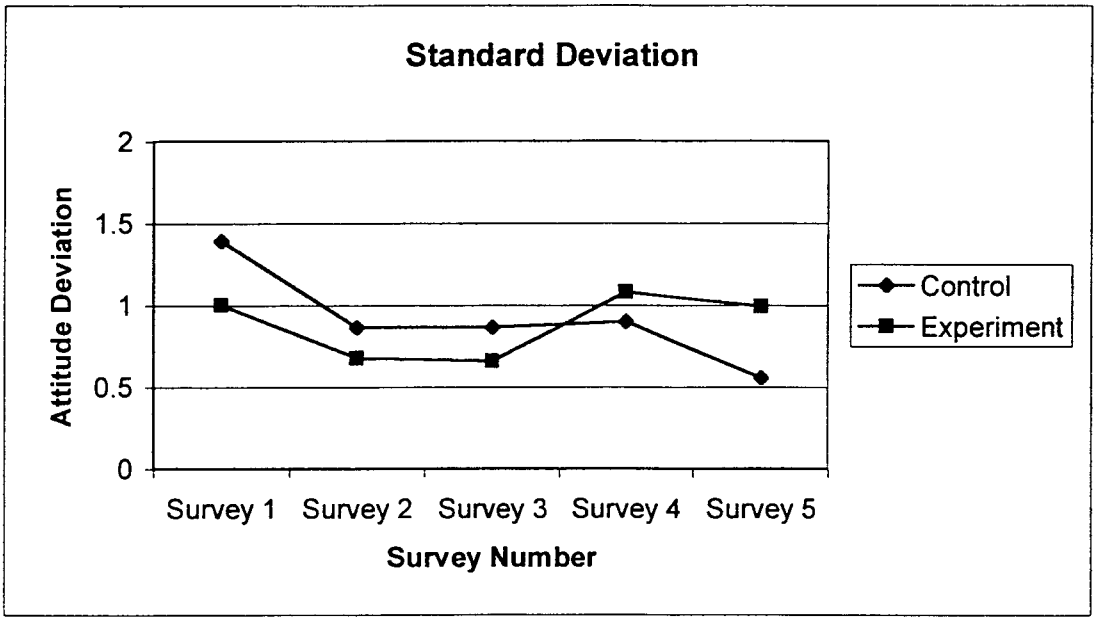
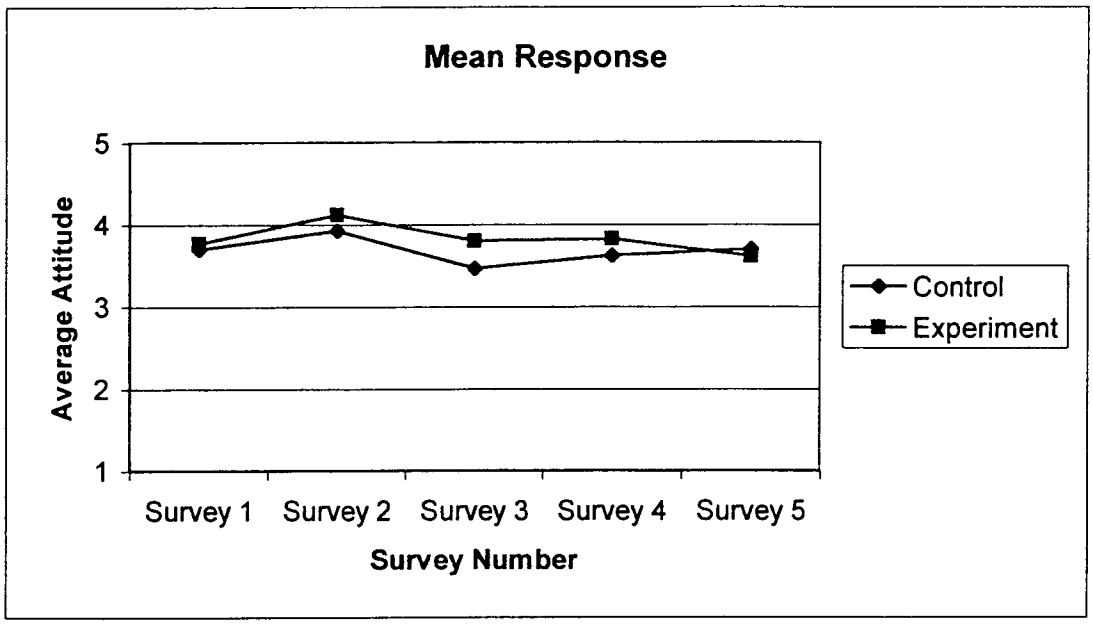


Figure C.1 Survey Question: How well satisfied are you with you current class standing?; a) mean response, b) response deviation

a)



b)

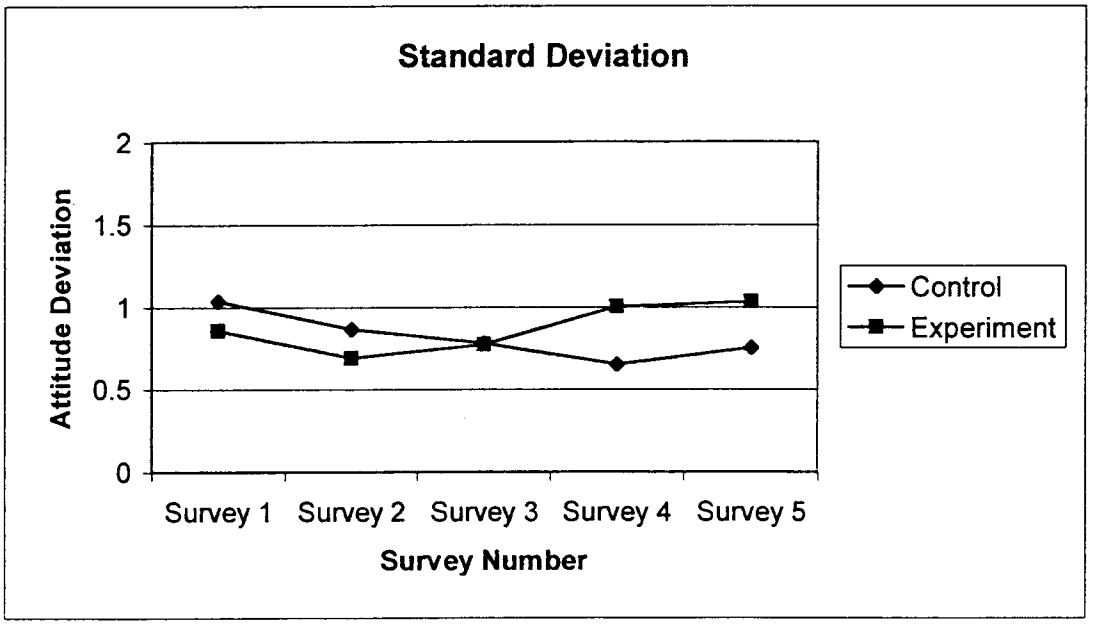
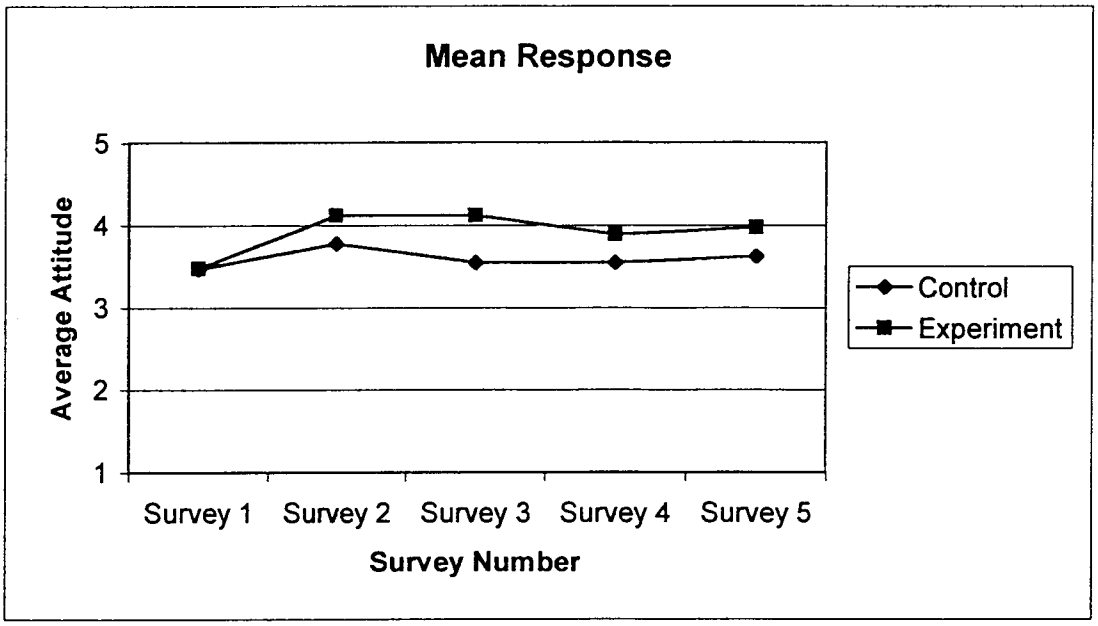


Figure C.2 Survey Question: How well satisfied are you with your chance of improving your evaluation in this course?; a) mean response, b) response

a)



b)

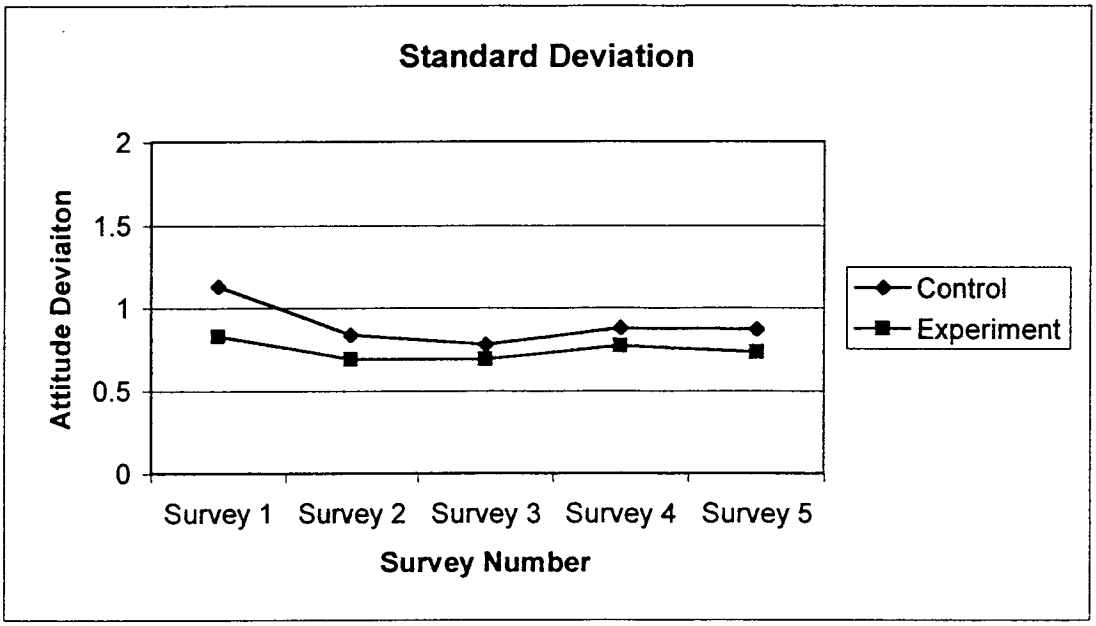
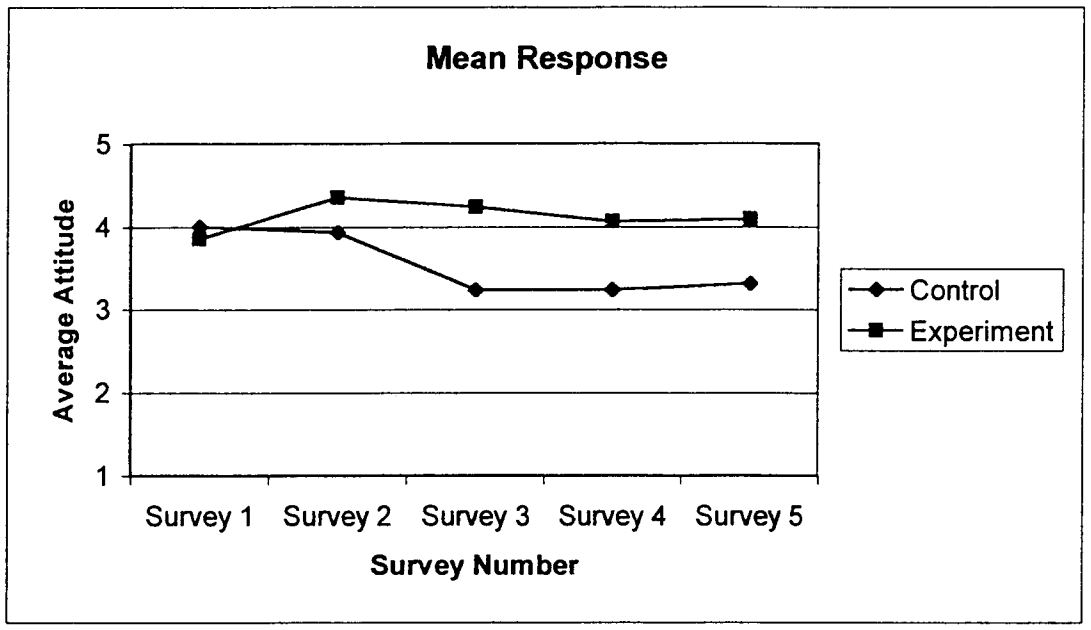


Figure C.3 Survey Question: How satisfied are you with the way things are going in this class?; a) mean response, b) response deviation

Appendix D: Satisfaction Results

a)



b)

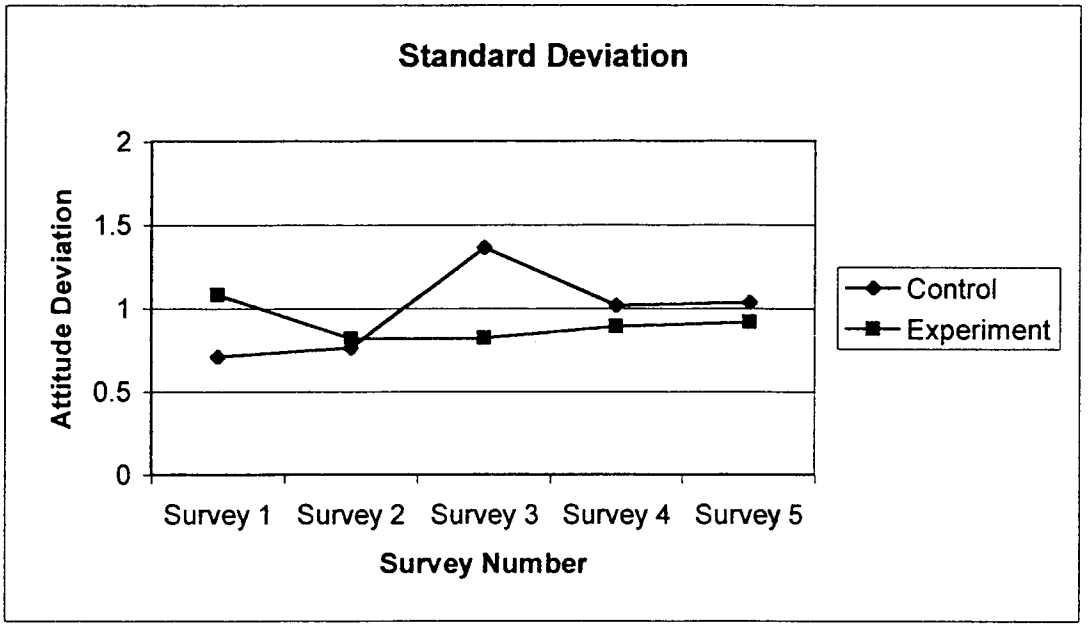
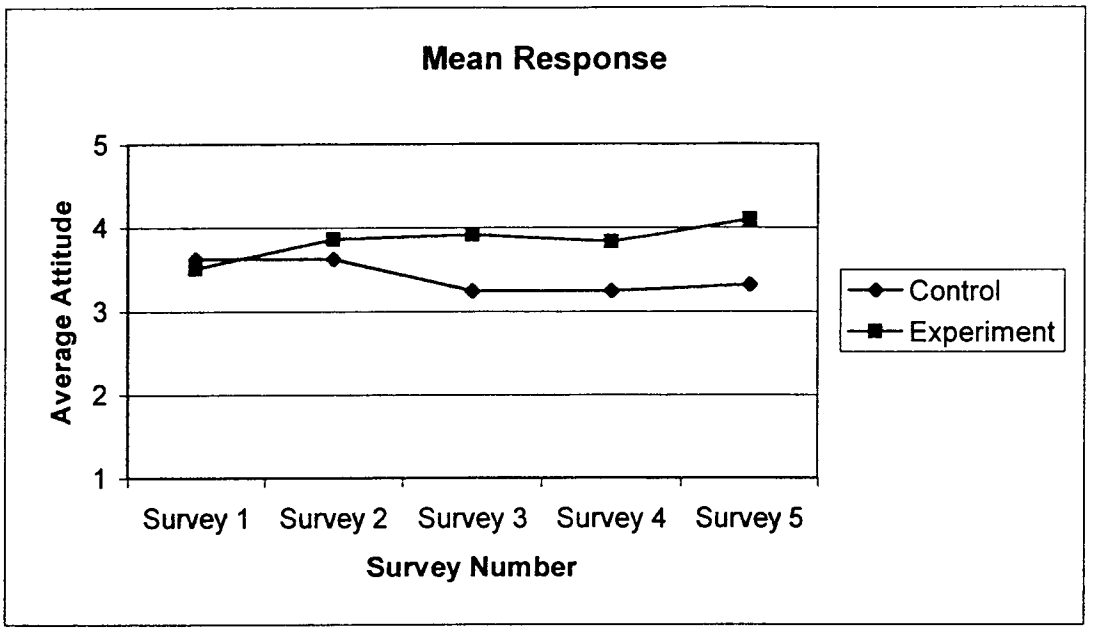


Figure D.1 Survey Question: How well did you like the sort of work you are doing?; a) mean response, b) response deviation

a)



b)

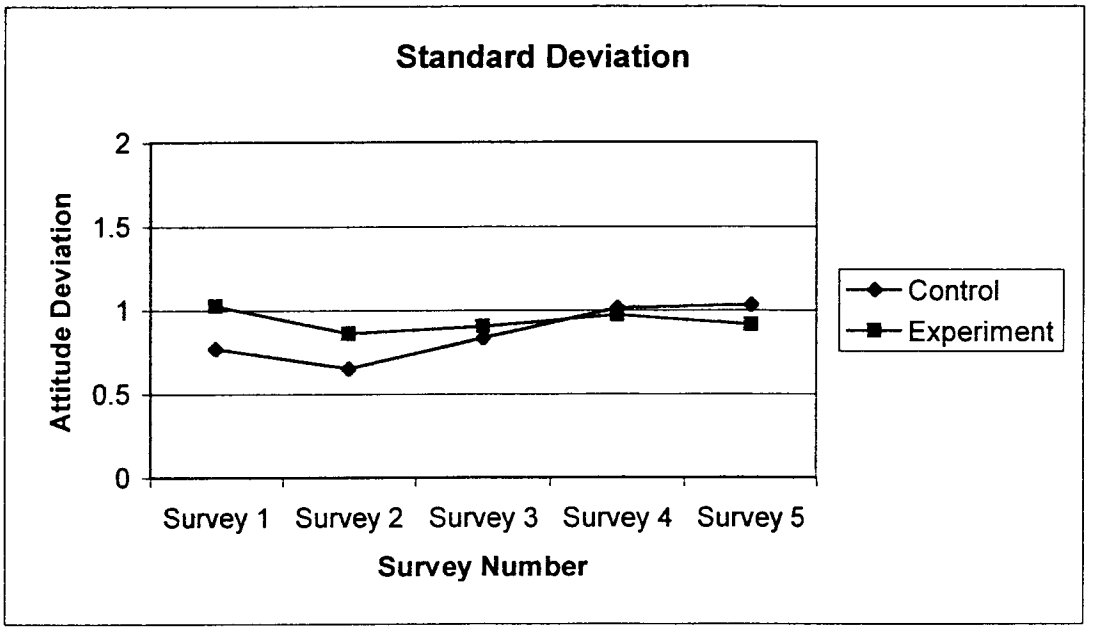
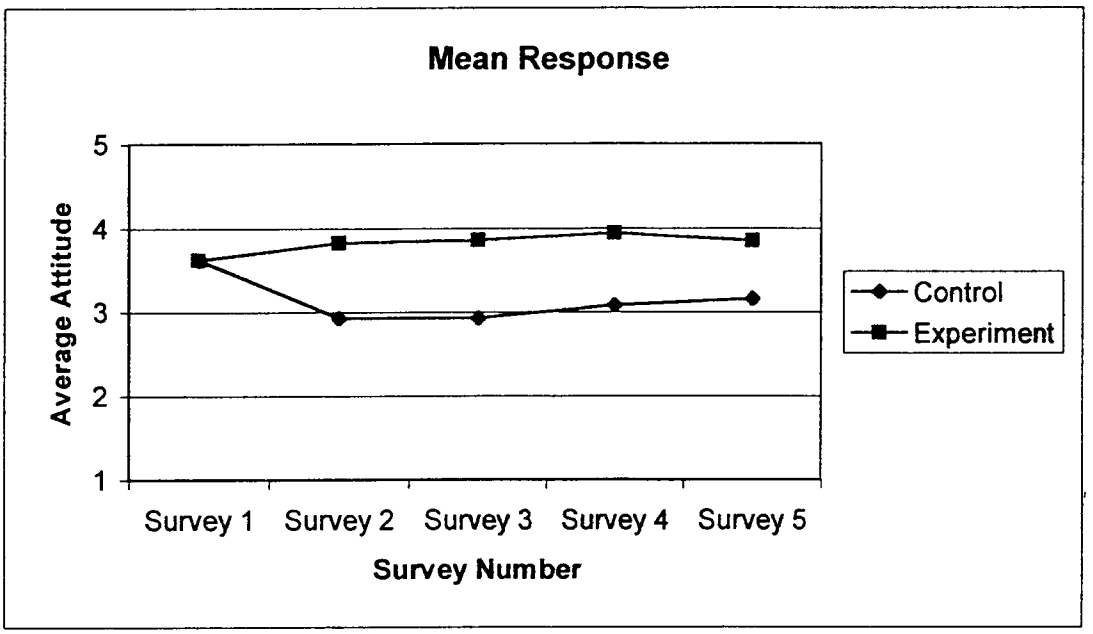


Figure D.2 Survey Question: Do you feel the assignments give you a chance to do what you do best?; a) mean response, b) response deviation

a)



b)

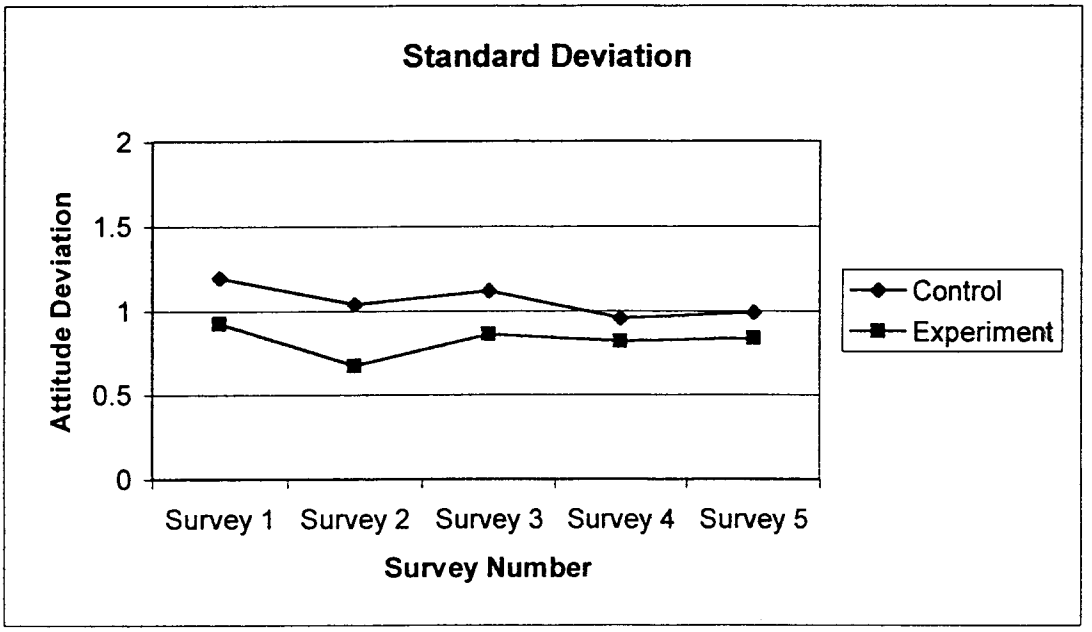
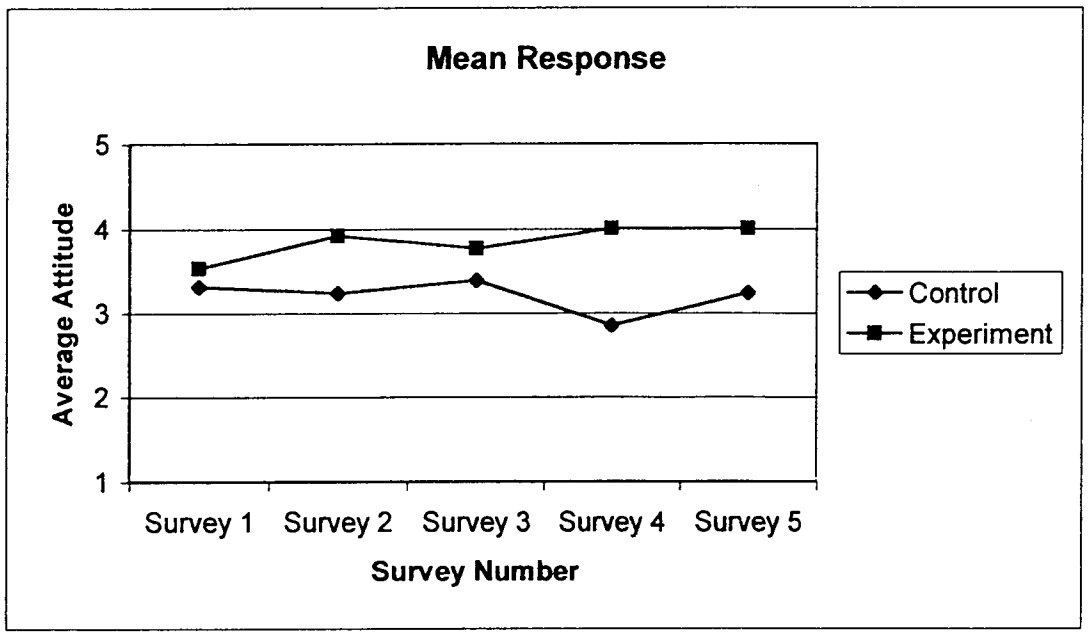


Figure D.3 Survey Question: Do you feel a sense of accomplishment from the work that you are doing?; a) mean response, b) response deviation

a)



b)

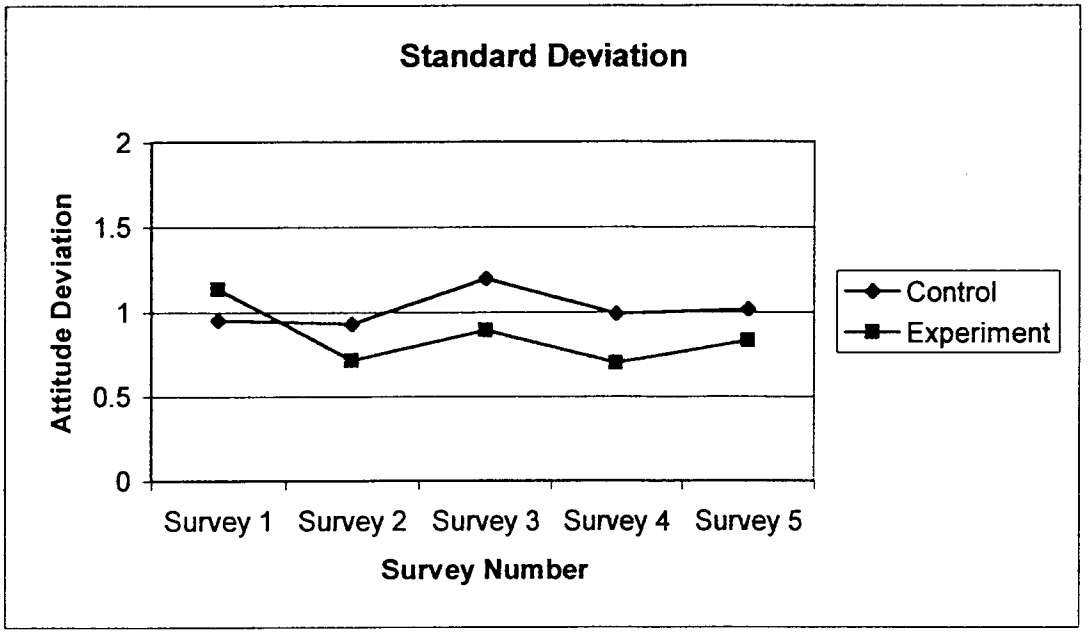
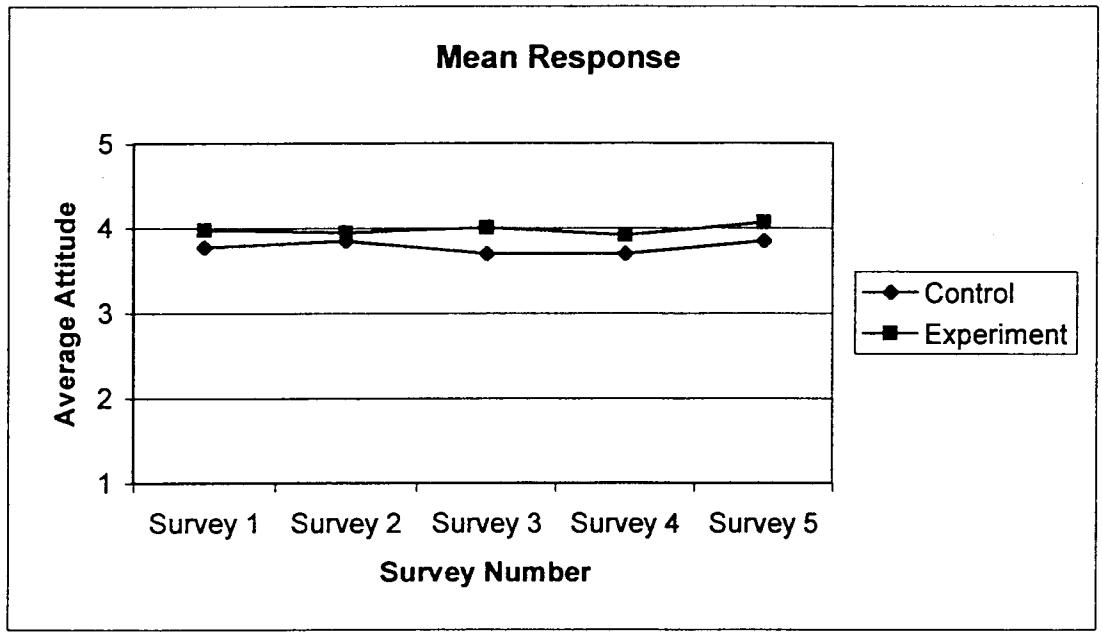


Figure D.4 Survey Question: I feel my contribution in this course is of (answer from scale of importance)?; a) mean response, b) response deviation

Appendix E: Individual Equity Results

a)



b)

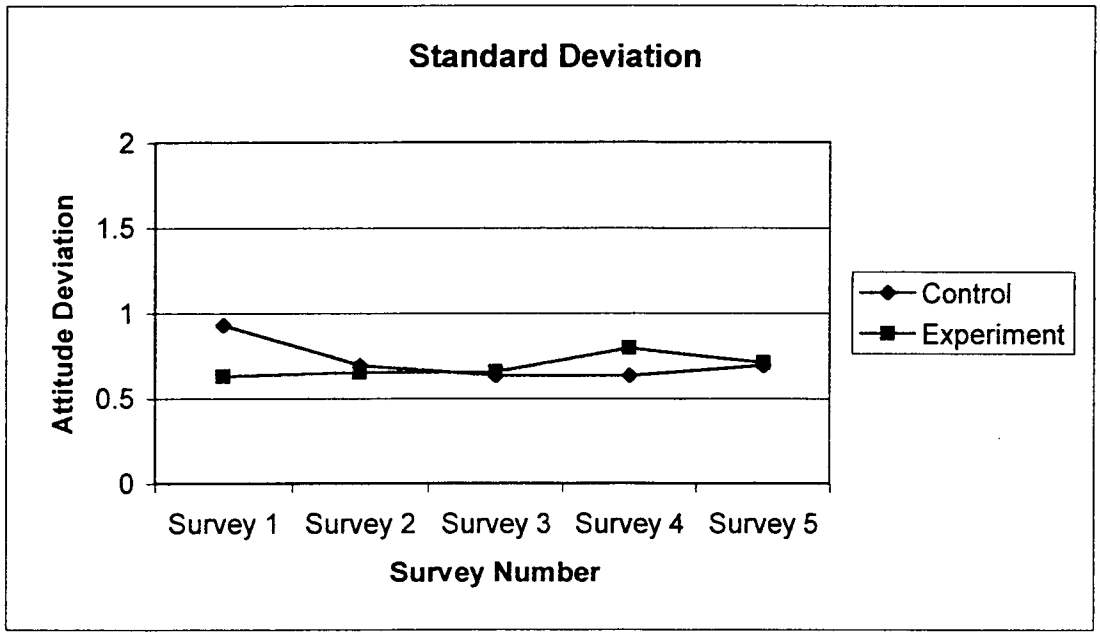
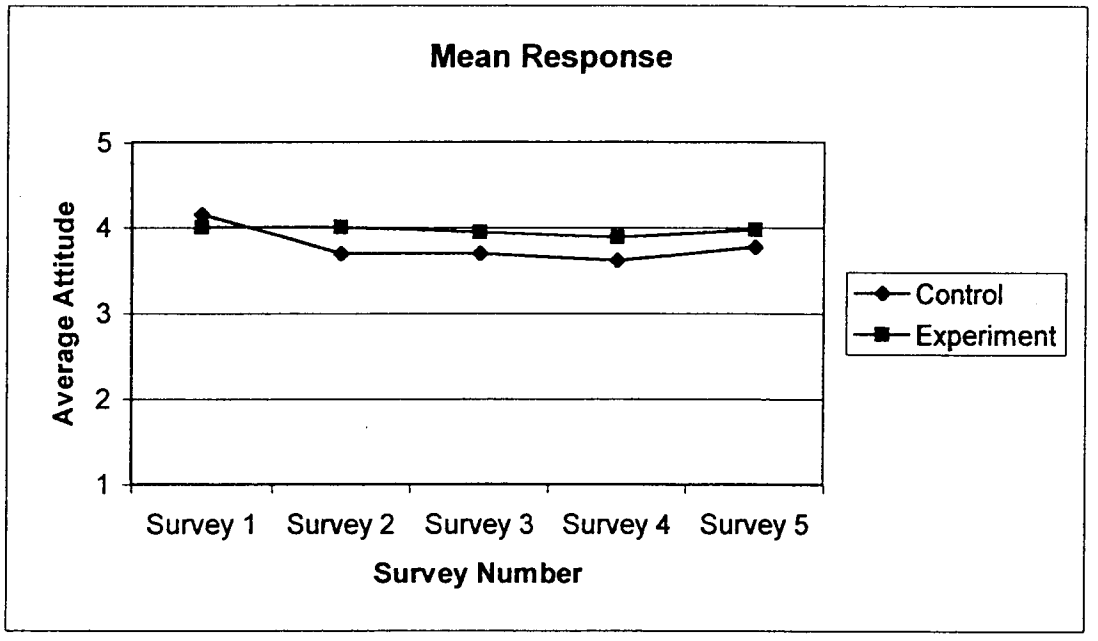


Figure E.1 Survey Question: How well do you think you compare with others in the class at getting things done?; a) mean response, b) response deviation

a)



b)

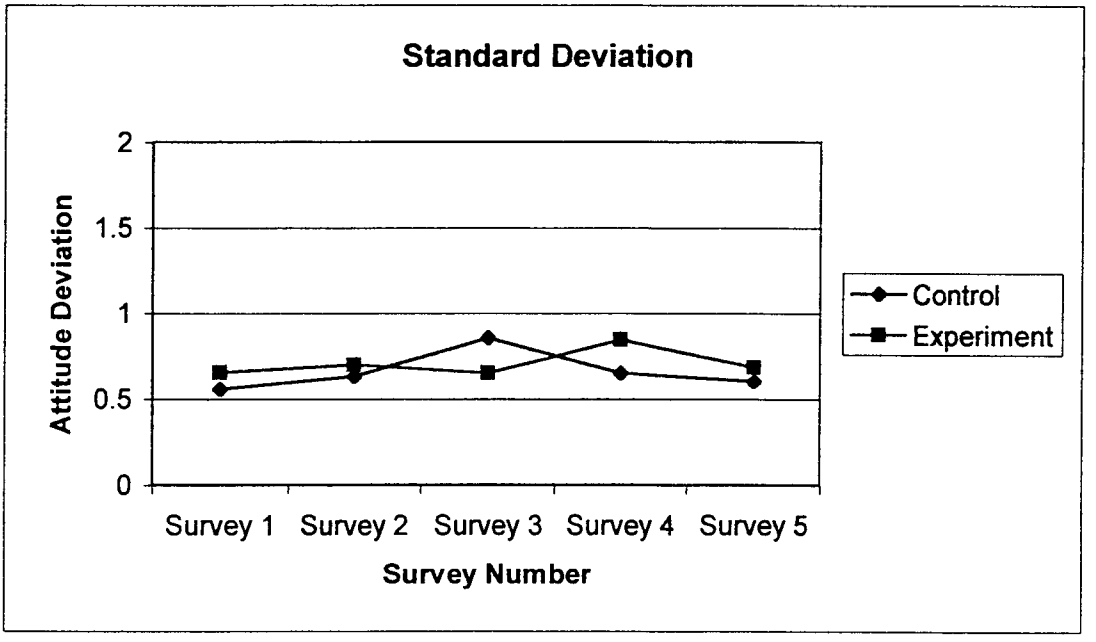
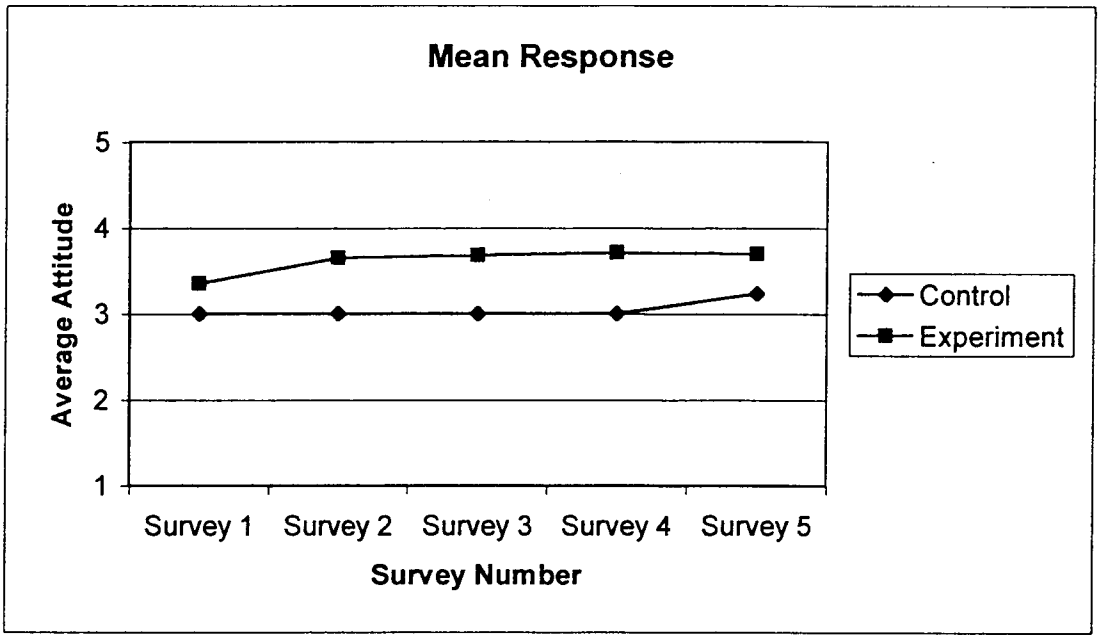


Figure E.2 Survey Question: How well do you think you compare with others in the class in quality of work?; a) mean response, b) response deviation

a)



b)

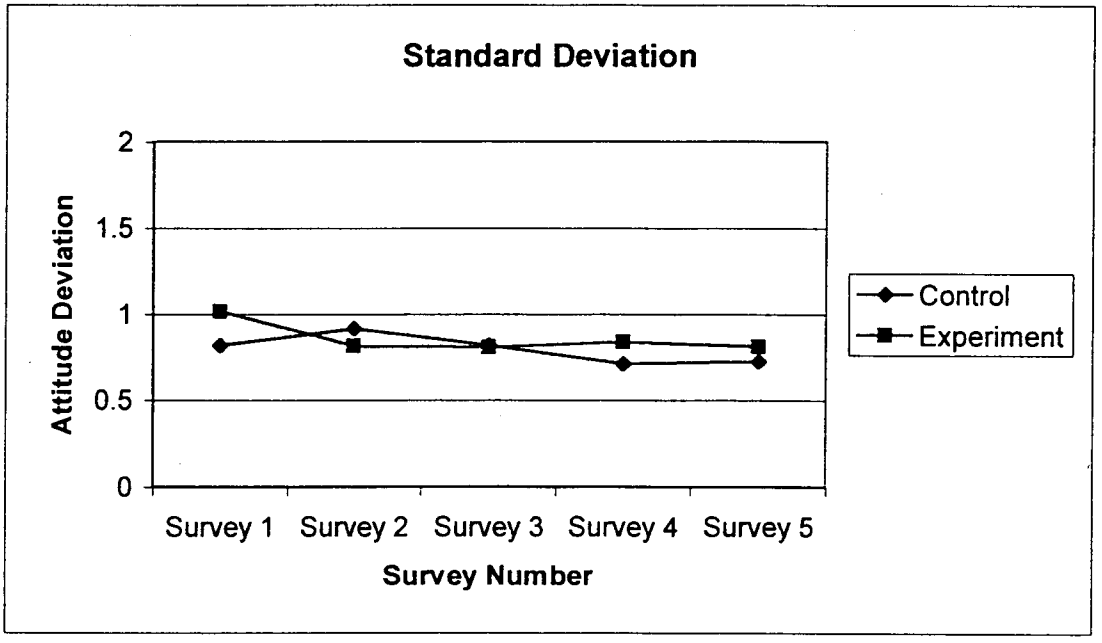
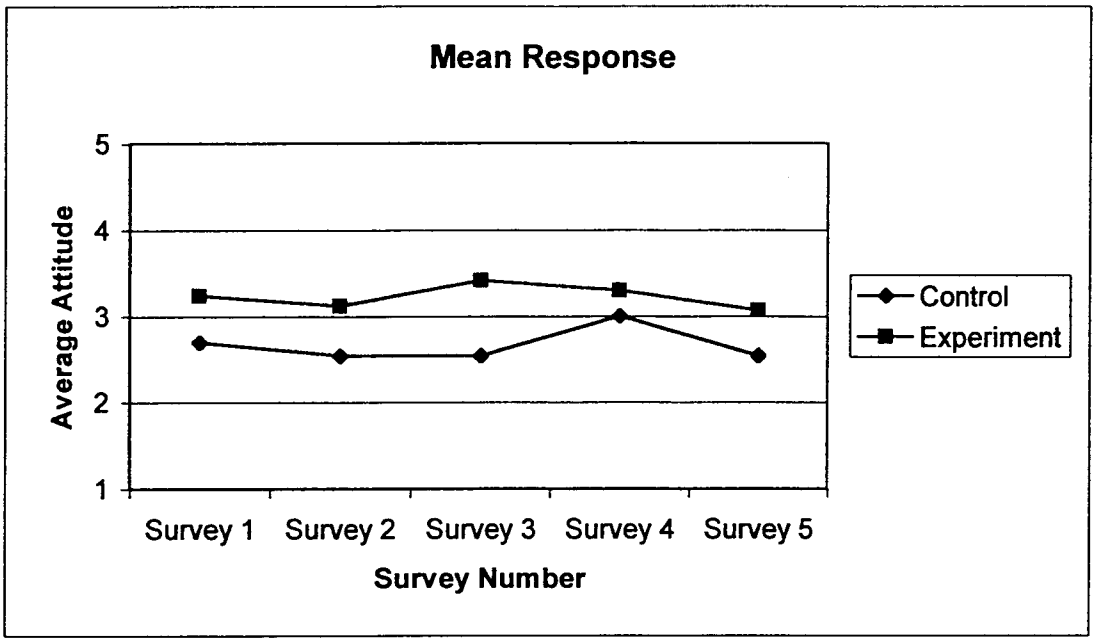


Figure E.3 Survey Questions: How well do you think you compare with others in the class in contribution to discussion?; a) mean response, b) response deviation

a)



b)

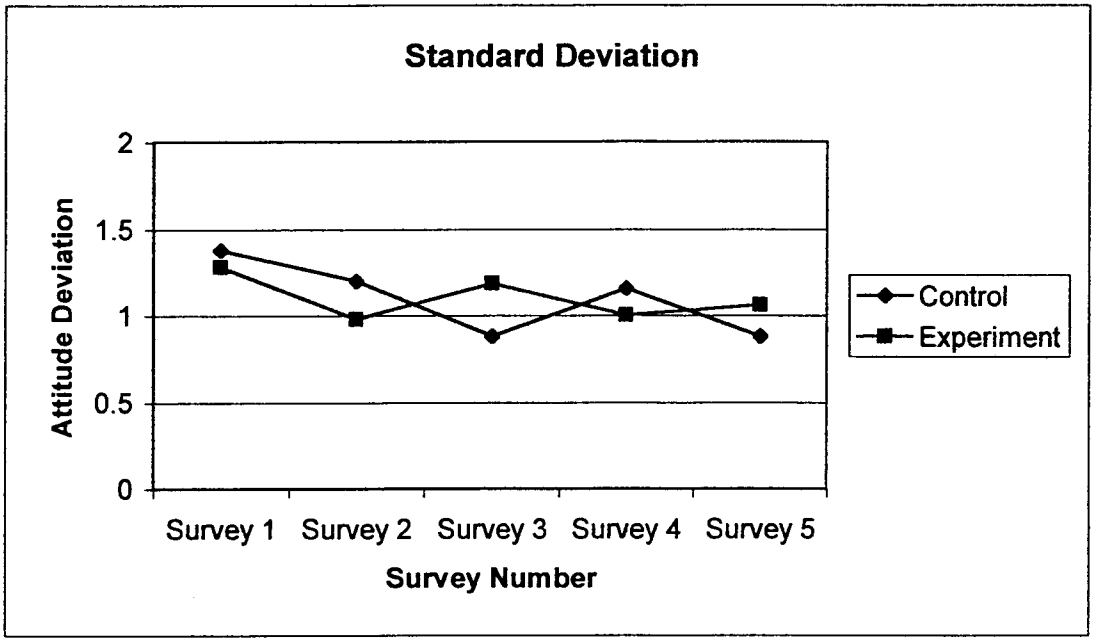


Figure E.4 Survey Question: What level of identification do you feel with the others in class?; a) mean response, b) response deviation

Appendix F: Focus Group Instructions and Questions

Focus Group Instructions

The purpose of this focus group is to better understand the relationship between the Balanced Scorecard Assessment used in IE 470/570, and the students in the course. The questions are open-ended, and there are no right or wrong answers. You will be given an opportunity to explain both positive and negative aspects of this assessment method. The intent is to find out how this assessment method effected the attitude of the students throughout the term. During the course of this focus group, responses will be written down to serve as a working memory for the group during this session, and as source of neutral information for analysis. In addition, responses will be tape recorded to serve as a backup during the analysis of the answers. The tape recording will only be reviewed by myself, and shared with no other individual. Your identity will be kept confidential, and your responses will only be used to support the research that has occurred over the course of this term.

Focus Group Questions:

1. What have you disliked about the balanced scorecard assessment method?

2. What have you liked about the balanced scorecard assessment method?

3. How do you feel the balanced scorecard compares with traditional assessment methods?

4. Would you want to take another course that utilized a balanced scorecard assessment method?

5. Did you feel that the assessment methods were fair?

6. How did the assessment measures in this class effect you motivation?

7. At any point during the term, were you frustrated with the way you were assessed in the class? Why?

8. Is there anything else that you would like to add about your experience with the assessment method this term?